

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
18. März 2004 (18.03.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/022780 A2

(51) Internationale Patentklassifikation⁷: C12Q 1/68

(21) Internationales Aktenzeichen: PCT/EP2003/008393

(22) Internationales Anmeldedatum:
30. Juli 2003 (30.07.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 38 434.7 16. August 2002 (16.08.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von
US): METANOMICS GMBH & CO. KGAA [DE/DE];
Tegeler Weg 33, 10589 Berlin (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): PLESCH, Gunnar
[DE/DE]; Karl-Marx-Str.43, 14482 Potsdam (DE). BLAU,
Astrid [DE/DE]; Rotkehlchenweg 33, 14532 Stahnsdorf
(DE). DÄSCHNER, Klaus [DE/DE]; Am Sandwerder 49,
14109 Berlin (DE).

(74) Anwalt: DÖRPER, Thomas; c/o BASF Aktienge-
sellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,
PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu ver-
öffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Ab-
kürzungen wird auf die Erklärungen ("Guidance Notes on Co-
des and Abbreviations") am Anfang jeder regulären Ausgabe der
PCT-Gazette verwiesen.

53851
021408

WO 2004/022780 A2

(54) Title: METHOD FOR IDENTIFYING SUBSTANCES HAVING A HERBICIDE ACTION

(54) Bezeichnung: VERFAHREN ZUR IDENTIFIZIERUNG VON SUBSTANZEN MIT HERBIZIDER WIRKUNG

(57) Abstract: The invention relates to a method for identifying compounds having a herbicide action. The invention also relates to nucleic acid constructs, to vectors containing said nucleic acid constructs, to transgenic organisms, and to the use of the same. Also disclosed are substances which have been identified by means of the abovementioned method.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein Verfahren zur Identifizierung von Verbindungen mit herbizider Wirkung. Weiterhin betrifft die Erfindung Nukleinsäurekonstrukte, Vektoren enthaltend die Nukleinsäurekonstrukte, transgene Organismen und deren Verwendung. Ausserdem betrifft die vorliegende Erfindung Substanzen, die mit dem vorgenannten Verfahren identifiziert wurden.

Verfahren zur Identifizierung von Substanzen mit herbizider Wirkung

Beschreibung

5 Die vorliegende Erfindung betrifft ein Verfahren zur Identifizierung von Verbindungen mit herbizider Wirkung. Weiterhin betrifft die Erfindung Nukleinsäurekonstrukte, Vektoren enthaltend die Nukleinsäurekonstrukte, transgene Organismen und deren Verwendung. Außerdem betrifft die vorliegende Erfindung Substanzen, die mit dem vorgenannten Verfahren identifiziert wurden.

10

Eine moderne Landwirtschaft ist ohne den Einsatz von Herbiziden nicht denkbar. Gegenwärtig wird der Wert der auf der gesamten Welt eingesetzten Herbizide auf ca. 30 Mrd. DM geschätzt. Obwohl bereits eine ganze Reihe von hoch wirksamen und ökologisch unbedenklichen Herbiziden auf dem Markt sind, ergibt sich die Notwendigkeit für neue Herbizide zum einen aus der Tatsache, dass immer wieder Unkräuter eine Resistenz gegen bereits eingesetzte Herbizide entwickeln und diese somit zum Teil nicht mehr eingesetzt werden können, zum anderen aus der Tatsache, dass ein Teil der Herbizide ökologische Nachteile aufweist. Zum heutigen Zeitpunkt werden in vielen Fällen noch Herbizide als Mischungen eingesetzt werden, die mehrere aktive Wirkstoffkomponenten enthalten, was ökologisch wenig vorteilhaft ist und zudem besondere Anforderungen an die Formulierung stellt.

15

20

25

Neue Herbizide sollten sich durch einen möglichst breiten Wirkungsbereich, ökologische und toxikologische Unbedenklichkeit sowie geringe Aufwandsmengen auszeichnen.

30

35

40

Das bisherige Vorgehen bei der Identifizierung und Entwicklung neuer Herbizide ist durch die Applikation von potentiellen Wirkstoffen direkt auf geeignete Testpflanzen gekennzeichnet. Dieses Verfahren weist den Nachteil auf, dass für die Tests relativ große Substanzmengen erforderlich sind. Dies ist im Zeitalter der kombinatorischen Chemie mit seiner in einer äußerst großen Vielfalt, aber nur in geringen Mengen herstellbaren Substanzen jedoch selten gegeben und stellt von daher eine wichtige Limitierung bei der Entwicklung neuer Herbizide dar. Zum anderen werden bei der direkten Applikation auf die zu testenden Pflanzen bereits im ersten Screeningschritt äußerst hohe Anforderungen an die Substanz gestellt, da nicht nur die Inhibierung oder sonstige Modulation der Aktivität eines zellulären Zieles (in der Regel ein Protein oder Enzym) erforderlich sind, sondern die Substanz dieses Ziel zunächst überhaupt erreichen muss, was bereits in diesem ersten Schritt Anforderungen an die Testsubstanz in Bezug auf Aufnahme durch die Pflanze, Permeabilität durch die verschiedenen Zellwände und Membranen, Persistenz zur Erreichung des gewünschten

Effektes, und schließlich Inhibierung/Veränderung der Aktivität des gewünschten Zielenzyms erfordert.

Es ist angesichts dieser Erfordernisse daher nicht überraschend, dass zum einen die Identifizierung neuer Wirkstoffe immer höhere Kosten verursacht, zum anderen immer weniger Wirkstoffe entdeckt werden.

Es war deshalb Aufgabe der vorliegenden Erfindung, Targets für die Identifizierung neuer Herbizide, sowie neue Herbizide und deren Verwendung zur Verfügung zu stellen. Diese Aufgabe wurde durch ein Verfahren zur Identifizierung von Substanzen mit herbizider Wirkung gelöst, dadurch gekennzeichnet, dass

a) die Expression oder die Aktivität des Genprodukts einer Nukleinsäure oder eines Gens umfassend:

15

aa) Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz;

20

25

bb) Nukleinsäuresequenz, die sich aufgrund des degenerierten genetischen Codes aus den durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen ableiten lässt;

30

35

cc) Nukleinsäuresequenz, die ein Derivat oder ein Fragment der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,

40

SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Nukleinsäuresequenzen ist, und mindestens 60 % Homologie auf Nukleinsäureebene aufweist;

5

dd) Nukleinsäuresequenz, die für Derivate oder Fragmente der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen codiert, die mindestens 50 % Homologie auf Aminosäureebene aufweisen;

10

15

ee) Nukleinsäuresequenz, die für ein Fragment oder ein Epitope eines Polypeptides codiert, das spezifisch an einem Antikörper bindet, wobei der Antikörper spezifisch an ein Polypeptid bindet, das von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz codiert wird;

20

25

ff) Nukleinsäuresequenz, die für ein Fragment einer in aa) dargestellten Nukleinsäure codiert und das eine "Translation Releasing Factor"-Aktivität, eine Kobalamin-Synthese-Aktivität, Arginyl-tRNA-Synthase-Aktivität, eine RNA-Helicase-Aktivität, eine GTP-Bindeprotein-Aktivität, eine Pseudouridylat-Synthase-Aktivität, eine Adenylatkinase-Aktivität, eine Preprotein-Translocase secA Precursorprotein-Aktivität, eine DCL-Protein-Aktivität, eine Arginin-tRNA-Ligase-Aktivität, eine plastidäre Gluthathion-Reduktase-Aktivität, eine Transkriptionsfaktor Sigma-Aktivität, eine Calmodulin-Aktivität, eine INT6-Aktivität, eine Helikase YGL150c-Aktivität, eine RNA-bindende Aktivität, eine Hitzeschock-Transkriptionsfaktor-Aktivität, eine chloroplastidäre DNA-Nukleoid bindende Aktivität oder eine Met2-Typ Cytosin DNA Methyltransferase-Aktivität hat; und/oder

30

35

40

gg) Nukleinsäuresequenz, die für Derivate der Polypeptide mit den in
SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8,
SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16,
SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24,
5 SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32,
SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40,
SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48,
SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen
codiert, die mindestens 20 % Homologie auf Aminosäureebene aufweist
10 und eine äquivalente biologische Aktivität besitzt; oder

b) die Expression oder Aktivität einer Aminosäuresequenz, die von einer Nukleinsäuresequenz nach aa) bis gg) codiert wird,
15 beeinflusst wird und solche Substanzen ausgewählt werden, die die Expression oder die Aktivität reduzieren oder blockieren.

Unter "Expression" wird die Neusynthese in vitro und in vivo von Nukleinsäuren und durch Nukleinsäuren codierten Proteinen verstanden, insbesondere die der obengenannten Nukleinsäure- und Aminosäuresequenzen. Der Begriff "Expression" umfasst
20 alle bis zum reifen Protein oder dessen Abbau führenden Biosyntheseschritte, z.B. Transkription, Translation, Modifizierung oder Prozessierung von Nukleinsäuren und/oder Proteinen, z.B. prae- oder posttranskriptionelle Prozessierungsschritte oder posttranslationale Modifikationen, z.B. Splicing, Editing, Polyadenylierung, Capping,
25 Modifikationen von Aminosäuren, z.B. Glykosylierung, Methylierung, Acetylierung, Bindung von Coenzymen, Phosphorylierung, Ubiquitination, Bindung von Fettsäuren, Signalpeptidprozessierung, etc.

Wobei im Sinne der Erfindung unter "Transkription" die RNA-Synthese mithilfe einer
30 RNA-Polymerase in 5' 3'-Richtung anhand einer DNA-Matrize zu verstehen ist.
Unter "Translation" ist die Biosynthese von Proteinen in vitro und in vivo zu verstehen.
Unter Genprodukt ist jedes Molekül und jede Substanz zu verstehen, die aufgrund der Expression, z.B. der Transkription oder Translation einer Nukleinsäure, z.B. einer DNA oder RNA, z.B. eines Gens, entsteht, wobei der Begriff auch die folgenden
35 Prozessierungsprodukte, wie z.B. nach Splicing oder Modifizierung, umfasst. So wird unter Genprodukt z.B. eine prozessierte RNA, z.B. eine katalytische RNA wie ein Ribozym, eine funktionelle RNA, wie tRNAs oder rRNAs, oder eine codierende RNA, wie mRNA, verstanden. Als Folge der Translation einer mRNA wird ein Protein synthetisiert, das ebenfalls als "Genprodukt" verstanden wird. Proteine können
40 während und nach der Translation verschiedenen Prozessierungsschritten unter-

worfen sein, wie unten beispielhaft aufgezählt. Unter "Aktivität des Genprodukts" ist die biologische Aktivität bzw. Funktion einer RNA oder eines Proteins, wie beispielsweise die enzymatische Aktivität, die Transporteraktivität, die regulatorische Aktivität, die Rezeptorbindungseigenschaft, die Fähigkeit, bestimmte Proteine, Nukleinsäuren oder Metaboliten zu binden, z.B. in Proteinkomplexen, das heißt z.B. die regulative Eigenschaft oder die Transporterfunktion des Proteins oder der RNA, um nur einige zu nennen, zu verstehen, wie sie im Organismus natürlicherweise vorkommt. Unter "Reduktion der Aktivität des Genprodukts" ist eine Verringerung der biologischen Aktivität im Vergleich zur natürlichen Aktivität des Genprodukts von mindestens 10 %, vorteilhaft mindestens 20 % oder 30 %, bevorzugt mindestens 40 %, 50 % oder 60 %, besonders bevorzugt um mindestens 70 %, 80 % oder 90 % und ganz besonders bevorzugt um mindestens 95 %, 96 %, 97 %, 98 % oder 99 % zu verstehen. Blockierung der Aktivität des Genprodukts bedeutet die gänzliche, das heißt 100 %ige Inhibierung der Aktivität oder die teilweise Blockierung der Aktivität, bevorzugt eine mindestens 80 %ige oder 90%ige, besonders bevorzugt mindestens 91%ige, 92%ige, 93%ige, 94%ige oder 95%ige, ganz besonders bevorzugt mindestens 95 %ige, 96%ige, 97%ige, 98%ige oder 99%ige Blockierung der biologischen Aktivität.

Eine Reduktion der Aktivität des Genprodukts kann auch indirekt erfolgen, z.B. in dem die Bildung oder Aktivität von Interaktionspartnern gehemmt wird, z.B. in dem die Stoffwechselkette, in die das Genprodukt eingebunden ist, beeinflusst wird. Z.B. kann eine Hemmung nicht nur des fraglichen Enzyms, sondern auch eines in derselben Stoffwechselkette stehenden Enzyms oder Proteins erfolgen, die zu einer Blockierung des folgenden, vorhergehenden oder eines anderen involvierten Enzyms führt und somit des hierin beschriebenen Genproduktes, z.B. durch Substrat- oder Produkt- hemmung. Solche Reduktionen durch indirekte Beeinflussung der Aktivität eines Enzyms ist z.B. für die Wechselwirkung der Glykolyseproteine und -metaboliten ausführlich beschrieben worden und leicht übertragbar auf andere Stoffwechselwege, in denen die hierin beschriebenen Genprodukte eine Rolle spielen. Genauso kann ein erfindungsgemäß verwendetes Genprodukt in seiner Aktivität reduziert oder gehemmt werden, in dem die Aktivität von Interaktionspartnern, z.B. anderen Proteinen in einem Proteinkomplex oder in einer Substrattransportkette mit dem hierin beschriebenen Genprodukt reduziert oder gehemmt wird. Das kann dazu führen, dass der gesamte Komplex oder der Substrattransport nicht mehr aktiviert oder nicht oder nur teilweise entsteht bzw. funktioniert oder nicht mehr regulierbar ist. Beispiele für solche Beeinflussung der Aktivität sind z.B. für Spleißosomen, Polymerasen, Ribosomen etc. beschrieben.

Unter "Fragment" wird eine Teilsequenz einer hierin beschriebenen Sequenz verstanden, die weniger Nukleotide oder Aminosäuren umfasst, als die hierin beschrie-

benen Sequenzen. Ein Fragment kann z.B. 1 %, 5 %, 10 %, 20 %, 50 %, 70 %, 90 % der ursprünglichen Sequenz umfassen. Vorzugsweise umfasst ein Fragment 100, mehr bevorzugt 50, noch mehr bevorzugt weniger als 20 Aminosäuren der entsprechenden Nukleinsäuren.

5

Die Bedeutung der einzelnen Biosyntheseschritten ist dem Fachmann bekannt und kann z.B. in "Molecular Biology of the cell", Alberts, New York, 1998, "Biochemie" Stryer, 1988, New York, "Biochemieatlas", Michal, Heidelberg, 1999 oder in "Dictionary of Biotechnology", Coombs, 1992, nachgelesen werden.

10

Eine Ausführungsform betrifft somit ein erfindungsgemäßes Verfahren, wobei die Expression oder die Aktivität der genannten Nukleinsäuren oder Aminosäuren dadurch reduziert oder blockiert wird, dass die Transkription, Translation, Prozessierung und/oder Modifikation mindestens einer der erfindungsgemäßen Nukleinsäuresequenz oder Aminosäuresequenz reduziert oder blockiert wird. Erfindungsgemäß können ein, zwei, drei oder mehr Sequenzen in ihrer Aktivität reduziert oder blockiert werden.

15

Das erfindungsgemäße Verfahren kann in einzelnen getrennten Verfahrensansätzen oder vorteilhaft in einem High-Throughput-Screening (HTS) durchgeführt werden und zur Identifizierung von Substanzen mit herbizider Wirkung oder von Antagonisten verwendet werden. Im vorgenannten Verfahren können vorteilhaft auch Substanzen identifiziert werden, die mit den oben genannten Nukleinsäuren bzw. mit deren Genprodukten interagieren, diese Substanzen sind potentielle Herbizide, die über die klassische chemische Synthese in ihrer Wirkung weiter verbessert werden können.

25

Nach dem Verfahren identifizierte bzw. ausgewählte Substanzen können vorteilhaft auf eine Pflanze verbracht werden, um die herbizide Aktivität der Substanzen zu testen. Es werden solche Substanzen ausgewählt, die eine herbizide Aktivität zeigen. In einer weiteren vorteilhaften Ausführungsform des Verfahrens können die Substanzen auch neben dem vorgenannten in vivo-Testverfahren auch in einem in vitro Test identifiziert werden. Ein derartiger in vitro Test mit den erfindungsgemäßen Nukleinsäuren bzw. deren Genprodukten hat den Vorteil, dass die Substanzen rasch und einfach auf ihre biologische Wirkung hin gescreent werden können. Derartige Test eignen sich auch vorteilhaft für das sog. HTS.

30

35

Das Verfahren kann mit freien Nukleinsäuren wie DNA oder RNA, freien Genprodukten oder vorteilhaft in einem Organismus durchgeführt werden, wobei als Organismus eukaryontische oder prokaryontische Organismen wie vorteilhaft gram-negative oder gram-positive Bakterien, Hefen, Pilze oder vorteilhaft Pflanzen wie monokotyle oder dikotyle Pflanzen verwendet werden. Als Organismen werden vorteilhaft konditionale

40

oder natürliche Mutanten verwendet, die die Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 betreffen.

Unter konditionalen Mutanten sind Mutanten zu verstehen, die erst nach Induktion eine Reduktion der Expression, z.B. der Transkription oder der Translation der vorher genannten Nukleinsäuren oder der durch sie codierten Genprodukte, aufweisen. Ein Beispiel derartiger konditionaler Mutanten sind Mutanten, in denen die Nukleinsäuren hinter einem temperatursensitiven Promotor liegen, der bei höheren Temperaturen nicht mehr funktionell ist, das heißt der die Transkription bei höheren Temperaturen beispielsweise oberhalb 37°C verhindert. Ebenfalls möglich ist beispielsweise eine Expressionsregulation durch ein Effektor-Molekül, z.B. bei Kontrolle der Expression durch einen regulierbaren Promotor, wie z.B. der im Tet-Systeme (Gatz et al., Plant J. 2, 1992:397-404, Tetracyclin-induzierbar) verwendete oder die in EP-A-0 388 186 (Benzylsulfonamid-induzierbar), EP-A-0 335 528 (Abzisin säure-induzierbar) oder WO 93/21334 (Ethanol- oder Cyclohexenol-induzierbar) beschriebenen Promotoren.

Eine weitere erfindungsgemäße Ausführungsform ist ein Verfahren zur Identifikation eines Antagonisten von Proteinen, die durch eine Nukleinsäuresequenz wie sie in dem erfindungsgemäßen Verfahren eingesetzt wird, insbesondere ausgewählt aus der Gruppe:

a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz,

- b) einer Nukleinsäuresequenz, die sich aufgrund des degenerierten genetischen Codes aus den durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen ableiten lässt,
- c) Nukleinsäuresequenz, die ein Derivat oder ein Fragment der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Nukleinsäuresequenzen ist, und mindestens 60 % Homologie auf Nukleinsäureebene aufweist,
- d) Nukleinsäuresequenz, die für Derivate oder Fragmente der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen codiert, die mindestens 50 % Homologie auf Aminosäureebene aufweisen;
- e) Nukleinsäuresequenz, die für ein Fragment oder ein Epitope eines Polypeptides codiert, das spezifisch an einem Antikörper bindet, wobei der Antikörper spezifisch an ein Polypeptid bindet, das der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz codiert wird;

- f) Nukleinsäuresequenz, die für ein Fragment einer in a) dargestellten Nukleinsäure codiert wird und das eine "Translation Releasing Factor"-Aktivität, eine Kobalamin-Synthese-Aktivität, Arginyl-tRNA-Synthase-Aktivität, eine RNA-Helicase-Aktivität, eine GTP-Bindeprotein-Aktivität, eine Pseudouridylat-Synthase-Aktivität, eine Adenylatkinase-Aktivität, eine Adenylatkinase-Aktivität, eine Preprotein-Translocase secA Precursorprotein-Aktivität, eine DCL-Protein-Aktivität, eine Arginin-tRNA-Ligase-Aktivität, eine plastidäre Gluthathion-Reduktase-Aktivität, eine Transkriptionsfaktor Sigma-Aktivität, eine Calmodulin-Aktivität, eine INT6-Aktivität, eine Helikase YGL150c-Aktivität, eine RNA-bindende Aktivität, eine Hitzeschock-Transkriptionsfaktor-Aktivität, eine chloroplastidäre DNA-Nukleoid bindende Aktivität oder eine Met2-Typ Cytosin DNA Methyltransferase-Aktivität hat; und/oder
- g) Nukleinsäuresequenz, die für Derivate der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen codiert, die mindestens 20 % Homologie auf Aminosäureebene aufweist und eine äquivalente biologische Aktivität besitzt; oder
- codiert werden, indem folgende Verfahrensschritte durchlaufen werden
- i) Inkontaktbringen von Zellen, die das Protein exprimieren, oder des Proteins mit einem Kandidatenstoff;
 - ii) Testen der biologischen Aktivität des Protein;
 - iii) Vergleichen der biologischen Aktivität des Proteins mit einer Standardaktivität in Abwesenheit des Kandidatenstoffs, wobei eine Verringerung der biologischen Aktivität des Proteins anzeigt, dass der Kandidatenstoff ein Antagonist ist.
- Unter ii) wird das Testen einer der oben beschriebenen biologischen Aktivitäten beschrieben, z.B. eine Enzymaktivität, wie sie in den Beispielen angegeben ist oder eine Bindung, vorzugsweise eine starke Bindung zwischen Protein- und Kandidatenstoff.

In einer vorteilhaften Ausführungsform des oben beschriebenen Verfahrens wird/werden der/die unter Buchstabe iii) identifizierte(n) Antagonist(en) auf eine Pflanze verbracht, um seine/ihre herbizide Aktivität zu testen und der/die Antagonist(en) ausgewählt, die eine herbizide Aktivität zeigen.

Das erfindungsgemäße Verfahren kann in einzelnen getrennten Verfahrensansätzen in vivo oder in vitro und/oder vorteilhaft gemeinsam oder besonders vorteilhaft in einem High-Throughput-Screening durchgeführt werden und zur Identifizierung von Substanzen mit herbizider Wirkung oder von Antagonisten verwendet werden.

Die im erfindungsgemäßen Verfahren identifizierten bzw. selektierten Nukleinsäuresequenzen sind für das Wachstum und die Entwicklung von höheren Pflanzen essentiell. Die Unterdrückung der Bildung der Genprodukte, d.h. der Expression, z.B. durch die spezifische Beeinflussung von z.B. Translation, Transkription oder Prozessierung und/oder der Unterdrückung der von den codierten Genprodukten ausgeübten Funktion bzw. biologischen Aktivität in intakten Pflanzen durch Substanzen vorteilhaft niedermolekularen Substanzen mit einem Molekulargewicht von kleiner 1000 Dalton, vorteilhaft kleiner 900 Dalton, bevorzugt kleiner 800, besonders bevorzugt kleiner 700, ganz besonders bevorzugt kleiner 600 Dalton, vorteilhaft mit einem Ki-Wert von kleiner 10^{-7} , vorteilhaft kleiner 10^{-8} , bevorzugt kleiner 10^{-9} M, vorteilhaft sollte dabei diese Hemmwirkung auf eine spezifische Hemmung der biologischen Aktivität der erfindungsgemäßen Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine zurückzuführen sein, das heißt es sollte keine Hemmung durch diese niedermolekularen Substanzen weiterer nahe verwandter Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine erfolgen. Weiterhin sollten die niedermolekulare Substanzen vorteilhaft ein Molekulargewicht von größer 50 Dalton, bevorzugt größer 100 Dalton, besonders bevorzugt größer 150 Dalton, ganz besonders bevorzugt größer 200 Dalton haben. Vorteilhaft sollten die niedermolekularen Substanzen weniger als drei Hydroxylgruppen an einem Kohlenstoffatom-enthaltenden Ring haben. Weiterhin sollten auch keine freie(n) Säure- oder Lacton-Gruppe(n) sowie keine Phosphatgruppe und nicht mehr als eine Aminogruppe im Molekül enthalten sein. Auch Basen wie Adenosin im Molekül sind weniger bevorzugt. Die über das erfindungsgemäße Verfahren identifizierten Substanzen vorteilhaft die niedermolekularen Substanzen, aber auch proteinogene Substanzen oder Sense- oder Antisense-RNA oder Antikörper oder Antikörperfragment führen durch ihre inhibitorischen Wirkungen vorteilhaft zu einer massiven Veränderung des Wachstums und der Entwicklung der behandelten bzw. betroffenen Pflanzen. Die im erfindungsgemäßen Verfahren identifizierten Substanzen sind deshalb in der Landwirtschaft als Herbizide geeignet.

Die in den erfindungsgemäßen Verfahren verwendeten Nukleinsäuren

SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 sind essentiell für Organismen, bevorzugt für Pflanzen. Ihre Unterbrechung bzw. die Blockierung ihrer Expression stoppt die Entwicklung von Pflanzen in einem frühen Entwicklungsstadium. Die Genprodukte der genannten Sequenzen sind z.B. den Polypeptiden der Sequenzen SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 zu entnehmen.

SEQ ID NO: 1, deren Expression in Linie 303317 blockiert ist, codiert für ein Protein (F2809.40), das Ähnlichkeiten zu translation releasing factor RF-2 aus *Synechocystis* sp. (PIR:S76448) aufweist und das auf Chromosom 3 (BAC ATF2809, Accession AL137080) von *Arabidopsis* liegt. Weiterhin besitzt das Protein die *araC*-Familien-signature.

SEQ ID NO: 3, deren Expression in Linie 304149 blockiert ist, codiert für ein "Kobalamin-Synthese"-Protein (MSH 12.9), das auf Chromosom 5 (P1-Klon MSH12, Accession AB006704) von *Arabidopsis* liegt.

SEQ ID NO: 5, deren Expression in Linie 120701 blockiert ist, codiert für einen ORF (T25K17.110) auf Chromosom 4 (BAC ATT25K17, Accession AL049171), der möglicherweise für eine Arginyl-tRNA-Synthetase codiert. Dieser ORF beinhaltet den EST: gb:AA404880, T76307.

SEQ ID NO: 7, deren Expression in Linie 126548 blockiert ist und auf Chromosom 4 des *Arabidopsis* Genoms (BAC ATF17A8, Accession AL049482) lokalisiert ist, codiert für ein putatives Protein (F17A8.80) mit Ähnlichkeit zu einer RNA-Helicase aus Maus (*Mus musculus*, PIR2:I84741).

Von SEQ ID NO: 9, deren Expression in Linie 127023 blockiert ist, wird ein putatives Protein (AT4g39780) codiert, das auf Chromosom 4 (BAC ATT19P19, Accession number AL022605) lokalisiert ist und das Homologien zu dem die AP2-Domäne

enthaltenden Protein RAP 2.4 aus *Arabidopsis thaliana* hat. Das Protein beinhaltet den ORF die ESTs gb146584 und AA394543.

5 SEQ ID NO: 11, deren Expression in Linie 127235 blockiert ist, codiert für den ORF F9K20.4, der auf Chromosom 1 (BAC F9K20, Accession AC005679) von *Arabidopsis* lokalisiert ist. Dieser ORF F9K20.4 codiert für ein putatives Protein mit Ähnlichkeit zu gil1786244 einem hypothetischen 24.9 kD Protein in der *surA-hepA* intergenischen Region *yab0* des *Escherichia coli* Genoms und zu gblAE000116 einem hypothetischen Protein der YABO Familie PF100849. Weiterhin besitzt das durch den ORF F9K20.4
10 codierte Protein eine konservierte Pseudouridylat-Synthase-Domäne, welche an der Modifizierung von Uracil in RNA-Molekülen beteiligt ist. Demnach zeigt der ORF F9K20.4 im blastp-Vergleich unter Standardbedingungen signifikante Homologie zu verschiedenen Pseudouridylat-Synthasen.

15 SEQ ID NO: 13, deren Expression in Linie 218031 blockiert ist, codiert für eine putative Adenylatkinase (At2g37250). Der ORF At2g37250 ist auf Chromosom 2 des Klons F3G5 (Accession AC005896) von *Arabidopsis* lokalisiert.

20 Das von SEQ ID NO: 15 codierte putative Protein (ORF T29H11_270, Accession AL049659), dessen Expression in Linie 171042 blockiert ist, zeigt Ähnlichkeit zu dem pol Polyprotein des "Equine Infectious Anemia Virus (PIR:GNLJEV). Die Sequenz befindet sich auf Chromosom 3 des BAC-Klons T29H11 von *Arabidopsis*.

25 SEQ ID NO: 17, dessen Expression in Linie KO_T3_02-33338-3 blockiert ist, befindet sich auf Chromosom 5 des P1-Klons MJE7 (Accession AB020745). Die Sequenz codiert für ORF MEJ7.11. Bei ORF MEJ7.11 handelt es sich um ein unbekanntes Protein.

30 SEQ ID NO: 19, dessen Expression in Linie KO_T3_02-33885-2 blockiert ist, codiert für ein unbekanntes Protein (= ORF F14G9.26). Der ORF befindet sich auf Chromosom 1 des BAC-Klons F14G8 mit der Accession AC069159.

35 SEQ ID NO: 21, dessen Expression in Linie KO_T3_02-35172-2 blockiert ist, codiert für ein unbekanntes Protein. Der ORF MAB16.6 hat nur Homologien zu anderen unbekannten Proteinen. Die Sequenz befindet sich auf Chromosom 5 des P1-Klons MAB16 mit der Accession AB018112.

40 SEQ ID NO: 23, dessen Expression in Linie 305861 blockiert ist, codiert für ein „Preprotein Translocase *secA* Precursor“ Protein, daher für ein chloroplastidäres *SecA*-Protein für den Proteintransport über die Thylakoidmembran. Dieser ORF findet

sich mit der Acc~~...~~on T7B11.6, AC007138, auf dem BAC-K~~...~~7B11 des Chromosoms 4.

5 Das von SEQ ID NO: 25 (= Linie 303814) codierte Protein mit der Accession F2G19.1, das signifikante Homologie zum DCL-Protein aus Tomate hat (PIR: S71749), findet sich auf dem BAC-Klon F2G19, Accession-Number AC083835, des Chromosoms 1.

10 SEQ ID NO: 27 (= Linie KO-T3-02-13224-1) kodiert für eine Arginin-tRNA-Ligase mit der Accession T25K17.110. Dieser ORF liegt auf dem BAC-Klon T25K17 mit der Accession-Number AL049171 und damit auf Chromosom 4.

15 SEQ ID NO: 29 (= Linie KO-T3-02-15114-2) codiert für eine plastidäre Guthathion-Reduktase. Dieser ORF mit der Accession T5N23.20 ist auf dem BAC-Klon T5N23, Accession-Number AL138650 auf Chromosom 3 annotiert.

20 SEQ ID NO: 31 (= Linie KO-T3-02-18601-1) codiert für ein Transkriptionsinitiationsfaktor Sigma Homologes. Dieser ORF mit der Accssion F22O13.2 ist auf dem BAC-Klon T22O13, Accession-Number AC003981, auf Chromosom 1 annotiert.

25 Durch SEQ ID NO: 33 (= Linie 304143) wird ein putatives Calmodulin-ähnliches Protein codiert. Dieser ORF mit der Accession At2g15680 ist auf dem BAC-Klons F9O13 mit der Accession-Number AC006248 auf Chromosom 2 annotiert.

30 Der von SEQ ID NO: 35 (= Linie KO-T3-02-40322-2) codierte unbekannte ORF MPX5.1 ist auf dem BAC-Klon MPX5, Accession-Number AP002048, auf Chromosom 3 annotiert.

35 SEQ ID NO: 37 (= Linie KO-T3-02-40309-1) codiert für ein Protein mit großer Ähnlichkeit zu INT6, einem brustkrebsassoziierten Protein und mit Ähnlichkeit zu einem „Initiationsfaktor 3“ Protein. Dieser ORF mit der Accession F28O9.140 ist auf dem BAC-Klon F28O9, Accession-Number AL137080, auf Chromosom 3 annotiert.

40 Das von SEQ ID NO: 39 (= Linie KO-T3-02-40309-1) codierte Protein hat große Ähnlichkeit zur DNA Helikase YGL150c aus Saccharomyces. Dieser ORF mit der Accesion F28O9.150 ist auf dem BAC-Klon F28O9, Accession-Number AL137080, auf Chromosom 3 lokalisiert.

45 SEQ ID NO: 41 (= Linie KO-T4-02-00666-4) codiert für ein Protein mit Ähnlichkeit zu einem RNA-bindenden Protein. Dieser ORF mit der Accession MKN22.2 befindet sich auf dem BAC-Klon MKN22, Accession-Nummer AB019234 des Chromosoms 5.

Von SEQ ID NO: 43 (= Linie KO-T4-02-00666-4) wird ein unbekanntes Protein codiert. Dieser ORF mit der Accession MEE6.19 ist auf dem BAC-Klon MEE6, Accession-Number AB010072, auf Chromosom 5 annotiert.

5

SEQ ID NO: 45 (= Linie KO-T3-02-41568-2) codiert für einen putativen Hitzeschock-Transkriptionsfaktor. Dieser ORF mit der Accession At2g26150 ist auf dem BAC-Klon T19L18, Accession-Number AC004747, auf Chromosom 2 lokalisiert..

- 10 Der in SEQ ID NO: 47 (= Linie KO-T3-02-42903-1) dargestellte ORF At2g28030 codiert für ein putatives chloroplastidäres an das DNA-Nukleoid bindende Protein. Dieser ORF At2g28030 ist auf dem BAC-Klon T1E2, Accession-Number AC006929, auf Chromosom 2 annotiert..

- 15 SEQ ID NO: 49 (= Linie KO-T3-02-41395-1) codiert für ein Protein mit Ähnlichkeit zu einer putativen Met2-Typ Cytosin DNA-Methyltransferase und weist starke Ähnlichkeit zu einer DNA-(cytosin-5-)-Methyltransferase aus *Arabidopsis thaliana* auf.. Dieser ORF mit der Accession AT4g08990 ist auf dem BAC-Klon ATCHRIV25, Accession-Number AL161513, auf Chromosom 4 annotiert.

20

Von SEQ ID NO: 51 (= Linie KO-T3-02-44634-4) wird ein Protein codiert, das große Ähnlichkeit zu einem postulierten Protein aus *Arabidopsis thaliana* hat. Dieser ORF mit der Accession F12B17_70 ist auf dem BAC-Klon F12B17, Accession-Number AL353995, auf Chromosom 5 lokalisiert.

- 25 Die genannten Sequenzen wurden alle in *Arabidopsis* identifiziert.

Die Unterdrückung der Bildung der Genprodukte bzw. Unterdrückung der von den codierten Genprodukten ausgeübten Funktion oder Aktivität in intakten Pflanzen durch eine niedermolekulare Substanz führt zur Reduzierung, bevorzugt zur Unterdrückung des Wachstums; die Entwicklung der Pflanze wird massiv verändert und unterdrückt. Sie sind deshalb zur Identifizierung von Herbiziden vorteilhaft geeignet.

30

Die vorgenannten Sequenzen oder funktionale Teile davon ermöglichen die Identifizierung von in der Landwirtschaft nutzbaren Herbiziden beispielsweise über ein Verfahren, das die folgenden Schritte umfasst:

35

- a) Bereitstellung zweier Linien eines Organismus, die die Genprodukte, die durch eine für das erfindungsgemäße Verfahren beschriebenen Sequenzen, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,

40

- SEQ ID NO: 7, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23,
SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31,
SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,
SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
5 SEQ ID NO: 49 oder SEQ ID NO: 51 codiert werden oder den beschriebenen
Derivaten oder Fragmenten davon, die die biologische Aktivität dieser
Sequenzen aufweisen, funktional exprimieren, wobei die Expression der Linien
unterschiedlich hoch ist, z.B. durch Mutagenese einer Linie und Identifizierung
10 einer Mutante mit erhöhter oder erniedrigter Expression und/oder Aktivität
des genannten Genprodukts im Vergleich zur Ausgangslinie oder, z.B. durch
Erzeugung von rekombinanten Organismen, vorteilhaft transgenen Pflanzen,
Pflanzengeweben wie Geweben von beispielsweise Blatt, Wurzel, Spross oder
Stamm, Pflanzensamen, Pflanzenkalli oder Pflanzenzellen, die die erfindungs-
gemäß beschriebenen Sequenzen, insbesondere SEQ ID NO: 1, SEQ ID NO: 3,
15 SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13,
SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID
NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31,
SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,
SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
20 SEQ ID NO: 49 oder SEQ ID NO: 51 oder Derivate oder Fragmente davon, die
die biologische Aktivität dieser Sequenzen besitzen, funktional exprimieren;
- b) Zugabe von chemischen Verbindungen (die auf ihre Herbizidaktivität getestet
werden sollen) zu den Linien mit den unterschiedlichen Expressions- oder Aktivi-
25 tätsleveln des Genprodukts, z.B. zu den unter a) genannten rekombinanten
Organismen und nicht-rekombinanten Ausgangsorganismen mit einem anderen,
vorzugsweise geringeren Expressions- oder Aktivitätslevel des Genprodukts;
- c) Bestimmung der biologischen Aktivität, beispielsweise der enzymatischen
30 Aktivität, des Wachstums oder der Vitalität der beiden Linien, z.B. der rekombi-
nanten Organismen, im Vergleich zu den nicht-rekombinanten Ausgangs-
organismen, nach Zugabe von chemischen Verbindungen gemäß Punkt b); und
- d) Selektion der chemischen Verbindungen, die die biologische Aktivität, beispiels-
35 weise die enzymatische Aktivität, das Wachstum oder die Vitalität der Linie mit
der geringeren Aktivität reduziert oder vollständig hemmt bzw. blockiert, z.B. die
die biologische Aktivität, das Wachstum oder die Vitalität der nicht rekombi-
nanten Organismen, der gemäß Punkt c) bestimmten chemischen Verbindun-
gen, im Vergleich zu den behandelten rekombinanten Organismen reduzieren
40 oder vollständig hemmen bzw. blockieren.

Ein in der Landwirtschaft nutzbares Herbizid kann auch identifiziert werden, wenn die oben in

- 5 a) erzeugten rekombinanten Organismen in einem Verfahren getestet werden, das folgende Schritte umfasst:
- (b) Zugabe von chemischen Verbindungen, die auf ihre Herbizidaktivität getestet werden sollen, zu den unter (a) genannten rekombinanten Organismen; und
- 10 (c) Bestimmung der biologischen Aktivität, beispielsweise der enzymatischen Aktivität, des Wachstums oder der Vitalität der rekombinanten Organismen nach Zugabe von chemischen Verbindungen gemäß (b) im Vergleich zu denselben nicht behandelten rekombinanten Organismen; und
- 15 (d) Selektion der chemischen Verbindung, die die biologische Aktivität, z.B. die enzymatische Aktivität, das Wachstum oder die Vitalität der behandelten Organismen im Vergleich zu den unbehandelten Organismen reduziert oder vollständig hemmt oder blockiert.

20 Unter chemischen Verbindungen, die die biologische Aktivität, das Wachstum oder die Vitalität der Organismen reduzieren, sind Verbindungen zu verstehen, die die biologische Aktivität, das Wachstum oder die Vitalität der Organismen mindestens um 10 %, 20 % oder 30 %, vorteilhaft um mindestens 40 %, 50 % oder 60 %, bevorzugt um mindestens 70 %, 80 oder 90 %, besonders bevorzugt um mindestens 91 %, 92 %, 25 93 %, 94 % oder 95 %, ganz besonders bevorzugt um mindestens 96 %, 97 %, 98 % oder 99 % hemmen, d.h. reduzieren oder blockieren.

Vorteilhaft ist insbesondere eine Substanz, die die Zelllinien mit geringer Aktivität 30 schädigen oder, vorzugsweise, die letal ist, jedoch nicht Zelllinien, die eine höhere Aktivität des Genproduktes aufweisen, schädigt oder für diese letal ist.

Allgemein können in dem genannten Verfahren Linien von Organismen eingesetzt werden, die die erfindungsgemäßen Sequenzen und insbesondere die Genprodukte, 35 die durch erfindungsgemäße Nukleinsäuren codiert werden, exprimieren, die jedoch nicht rekombinant sind, solange eine Linie eine höhere Genexpression oder Aktivität des Genprodukts aufweist als eine andere Linie. Solche Linien können natürlich auftreten oder durch Mutagenesen erzeugt werden.

Assaysysteme, die die Identifizierung von Substanzen, die die Bildung der Genprodukte und/oder die von den Genprodukten ausgeübten Funktionen oder die Aktivität der Genprodukte in intakten Pflanzen, Pflanzenteilen, -geweben oder Pflanzenzellen unterdrücken, sind dem Fachmann bekannt. Beispielfhaft sei hier auf Testsysteme für die Inhibierung von Enzymen wie der Adenylat-Kinase wie von Skoblov et al. (FEBS Letters, 395 (2-3), 1996: 283-285), von Russel et al. (J. Enzyme Inhib., 9 (3), 1995: 179-194 und), Wiesmüller et al. (FEBS Letters, 363, 1995: 22-24) oder Schlattner et al. (Phytochemistry, 42, 1996: 589-594) beschrieben, verwiesen. Diese Testsysteme können beispielsweise vorteilhaft für sog. Inhibierungsassays für beispielsweise das in Linie 218031 identifizierte Genprodukt verwenden.

Weitere vorteilhafte Assaysysteme sind beispielsweise die Fluoreszenz-Korrelations-spektroskopie (= FCS). Mit Hilfe der FCS (Brock et al., PNAS, 1999, 96, 10123-10128; Lamb et al., J. Phys. Org. Chem., 2000, 13654-658) ist es möglich, die zeitliche Diffusion von Molekülen zu messen bzw. die Differenz der gebundenen gegenüber den freien Molekülen zu ermitteln. Hierzu werden die zu untersuchenden Moleküle fluoreszenzmarkiert und beispielsweise ein definiertes Volumen in Mikrotiterplatten gegeben. Die Fluktuation der Moleküle wird in den Proben dabei von der Braunschen Molekularbewegung getrieben. Durch einen in der Probe fokussierten Laser lassen sich die translateralen bzw. rotationale Diffusion und Konformationsänderungen der Moleküle verfolgen und über eine Korrelation analysieren. Durch Bindung an andere Substanzen ändert sich der Diffusionskoeffizient der Moleküle. Mit Hilfe verschiedener Algorithmen lässt sich die Bindung der Moleküle über die Änderung des Diffusionskoeffizienten ermitteln bzw. quantifizieren. Mit dieser Methode kann in einem breiten Konzentrationsbereich vorteilhaft gemessen werden. Die Methode eignet sich vorteilhaft für die Messung von rekombinanten Proteinen, die vorteilhaft mit einem sog. His-Tag zur leichteren Aufreinigung über handelsübliche Chromatographie-Säulen (Porath et al., Nature 1975, 258, 598-599) versehen sind. Das so gereinigte Protein wird schließlich mit einem Fluoreszenzmarker wie z.B. Carboxytetramethylrhodamin oder BODIPY® (z.B. BODIPY 576/589 Angiotensin II, NEN® Life Science Products, Boston, MA, USA) versehen. Die zu untersuchende Verbindung bzw. Substanz wird zu dem Protein anschließend in einem Überschuss zugegeben. Die Diffusion des so markierten Proteins wird schließlich mit einem FCS-System (z.B. ConfoCor2 mit LSM 510, Carl Zeiss Mikroskop, Jena, Deutschland) ermittelt.

Eine weitere vorteilhafte Detektions-Methode für das erfindungsgemäße Verfahren ist die sog. "Surface Enhanced Laser Desorption Ionisation"-Methode (= SELDI Protein-Chip®). Diese Methode wurde von Hutchens und Yip (1980) erstmals beschrieben. Mit dieser Methode, die für die reproduzierbare, gleichzeitige Identifizierung von Biomarkern oder Antigenen entwickelt wurden (Hutchens und Yip, Rapid Commun. Mass

Spectrom, 1993 (576-580), kann die Ligand-Protein-Bindung über Massenspektrometrie analysiert werden. Dabei erfolgt die Detektion über normale TOF-Detektion (= time of flight). Auch bei dieser Methode können rekombinant exprimierte Proteine wie oben beschrieben exprimiert und gereinigt werden. Zur Messung wird das Protein auf den SELDI ProteinChips® immobilisiert, beispielsweise über die schon zur Reinigung verwendeten His-Tags oder über Ionen- oder hydrophobe Wechselwirkungen mit dem Chip. Auf diesen so vorbereiteten Chip werden anschließend die Liganden mit beispielsweise einem Autosampler gegeben. Nach einem oder mehreren Waschschritten mit Puffern verschiedener Ionenstärke werden die gebundenen Liganden mit dem LDI-Laser analysiert. Dabei wird die Bindungsstärke der Liganden nach jedem Waschschriff ermittelt.

Als weitere vorteilhafte Detektionsmethode sei die sog. Biacore-Methode genannt, bei der der Refraktionsindex an der Oberfläche bei Bindung von Liganden und das an der Oberfläche gebundene Protein analysiert wird. Bei dieser Methode wird eine Kollektion von kleinen Liganden sequentiell in eine Messzelle mit dem gebundenen Protein gegeben. Die Bindung an der Oberfläche wird über eine erhöhte sog. "Plasmon-Resonanz" (= SPR) über die Aufzeichnung der Laserrefraktion von der Oberfläche ermittelt. Im allgemeinen ist die Refraktionsindexänderung, die für eine Änderung der Massenkonzentration an der Oberfläche, ermittelt wird für alle Proteine oder Polypeptide gleich, das heißt diese Methode kann vorteilhaft für die verschiedensten Proteine verwendet werden. (Liedberg et al., Sens. Actuators, 1984, 4, 299-304). Wie oben beschrieben werden auch hier vorteilhaft rekombinant exprimierte Proteine verwendet, die an den Biacore Chip (Upsala, Schweden) beispielsweise über Histidin-Reste (z.B. His-Tag) gebunden werden. Der so hergestellte Chip wird wieder mit den Liganden in Verbindung gebracht z.B. mit einem Autosampler und die Bindung über ein von Biacore vertriebenes Detektionssystem mit Hilfe des SPR-Signals d.h. über die Änderung des Refraktionsindex gemessen.

Die erfindungsgemäßen Verfahren haben eine Reihe von Vorteilen wie beispielsweise:

- * neue potentielle Angriffsorte für herbizide Wirkstoffe können identifiziert werden,
- * Identifizierung von Herbiziden, die eine möglichst umfassende, Pflanzenspezies unabhängige Wirkung haben,
- * Substanzen, die mittels der kombinatorischen Chemie erzeugt wurden, und die sich durch eine hohe Vielfalt, aber geringe zur Verfügung stehende Mengen auszeichnen, können effizient auf Inhibitoren der neu identifizierten Angriffsorte geprüft werden

* sie erlauben, landwirtschaftlichen Nutzpflanzen im Fall von z.B. sehr breite
Wirksamkeit aufweisenden Herbiziden (Totalherbiziden oder auch selektiven
Herbiziden) Resistenz gegenüber diesen zu vermitteln (siehe Beschreibung
5 im folgenden)

Z.B. können Substanzen, die besonders spezifisch mit z.B. einem Protein oder
Proteinfragment binden, das von einer Nukleinsäure codiert wird, deren Expression
essentiell für das Wachstum der Pflanzen ist, mit den genannten Verfahren isoliert
10 werden. Dies ermöglicht eine vereinfachte Identifikation möglicher Inhibitoren, die
Proteine, z.B. in ihren Enzymeigenschaften, Bindeeigenschaften oder sonstigen
Aktivitäten hemmen, z.B. auch durch die Inhibierung ihrer Prozessierung, wie oben
beschrieben, oder ihren Transport innerhalb der Zelle oder Im- und Export aus
Organellen oder Zellen verhindern. Die so identifizierten Substanzen können auch in
15 einem weiteren Schritt in Screening-Verfahren, wie sie dem Fachmann bekannt sind,
auf Pflanzen aufgebracht werden und auf ihre Beeinflussung des Wachstums und der
Entwicklung hin untersucht werden. Somit wird eine Auswahl aus der unendlichen
Zahl chemischen Verbindungen, die sich für ein Screeningverfahren eignen würden,
getroffen, die es dem Fachmann wesentlich erleichtert, herbizide Substanzen zu
20 identifizieren.

Unter "spezifische Bindung" versteht man die Spezifität von Interaktionen zwischen
zwei Partnern, z.B. Proteinen untereinander oder von Protein (Enzym) und Substrat
(Substratspezifität). Sie beruht auf einer bestimmten molekularen räumlichen Struktur.
25 Wird sie zerstört, spricht man von Denaturierung, die oftmals irreversibel ist und
wodurch die Spezifität meistens verloren gehen kann. Diese biologische Aktivität ist
stark abhängig von den Umgebungsbedingungen (Puffer, Temperatur, Kontakte
zu unphysiologische Oberflächen wie Glas oder fehlende Cofaktoren). Bei Enzym-
Substrat oder Cofaktor, bei Rezeptor-Ligand oder bei Antikörper-Antigen Bindungen
30 spricht man von spezifischen Bindungen. Die Enzym-Substrat Wechselwirkung wird
thermodynamisch im einfachsten Fall mit der Michaelis-Menten-Gleichung beschrie-
ben. Sie beschreibt die Enzymaktivität über die sog. Michaelis-Menten-Konstante, die
wiederum die Kinetik widerspiegelt. Diese Konstante ist auch die Maßeinheit für die
Enzymaktivität, die wiederum die Spezifität widerspiegelt. Definition der Enzym-
35 aktivitätseinheit (nach IUB): Eine Einheit U entspricht derjenigen Enzymmenge, welche
die Umsetzung von einem Mikromol Substrat pro Minute unter genau festgelegten
Versuchsbedingungen katalysiert. Die spezifische Aktivität wird meist in U/mg ange-
geben.

In einem weiteren Schritt können dann die identifizierten Substanzen auf Pflanzen, Mikroorganismen oder Zellen aufgebracht werden, z.B. auf Pflanzenzellen, und dann die Beeinflussung des Stoffwechsels dieser Pflanzen beobachtet werden, z.B. Enzymaktivitäten, Photosyntheseaktivitäten, Stoffwechselaktivität, Fixierungsrate, Gasaustausch, DNA-Synthese, Wachstumsraten. Diese und viele andere dem Fachmann bekannte Methoden eignen sich, um die Viabilität von Zellen zu untersuchen. Substanzen, die das Wachstum, z.B. von Zellen, insbesondere Pflanzenzellen, reduzieren, insbesondere blockieren, eignen sich dann bevorzugt als Auswahl für herbizide Mittel.

10 Weiterhin können schon in einem sehr frühen Stadium Untersuchungen zu den Aufwandmengen der gefundenen Herbizide gemacht werden. Außerdem kann die hohe Spezifität und Effizienz gegenüber Unkräutern leicht ermittelt werden.

Mit Hilfe des erfindungsgemäßen Verfahrens kann eine Vielzahl von chemischen Verbindungen schnell und einfach auf herbizide Eigenschaften überprüft werden. Das Verfahren gestattet es, reproduzierbar aus einer großen Anzahl von Substanzen gezielt solche mit großer Wirkstärke auszuwählen, um mit diesen Substanzen anschließend weitere, dem Fachmann geläufige vertiefte Prüfungen durchzuführen.

20 Weiterer Gegenstand der Erfindung ist ein Verfahren zur Identifizierung von Inhibitoren pflanzlicher Proteine, die durch die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen codiert werden, mit potentiell herbizider Wirkung indem man die Genprodukte kloniert, in einer geeigneten Expressionskassette - beispielsweise in Insektenzellen - zur Überexpression bringt, die Zellen öffnet und den Zellextrakt
25 direkt bzw. nach Anreicherung oder Isolierung des Proteins in einem Testsystem zur Messung der biologischen Aktivität in Gegenwart von niedermolekularen chemischen Verbindungen einsetzt.

Ein weiterer Gegenstand der Erfindung sind deshalb Substanzen identifiziert nach den
30 erfindungsgemäßen Verfahren, wobei die Substanzen vorteilhaft niedermolekulare Substanzen mit einem Molekulargewicht von kleiner 1000 Dalton, vorteilhaft kleiner 900 Dalton, bevorzugt kleiner 800, besonders bevorzugt kleiner 700, ganz besonders bevorzugt kleiner 600 Dalton, vorteilhaft mit einem Ki-Wert von kleiner 10^{-7} , vorteilhaft kleiner 10^{-8} , bevorzugt kleiner 10^{-9} M sind, vorteilhaft sollte dabei diese Hemmwirkung
35 auf eine spezifische Hemmung der biologischen Aktivität der erfindungsgemäßen Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine zurückzuführen sein, das heißt es sollte keine Hemmung durch diese niedermolekularen Substanzen weiterer nahe verwandter Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine erfolgen. Weiterhin sollten die bevorzugten niedermolekularen Substanzen vorteilhaft ein Molekulargewicht von größer 50 Dalton,
40

bevorzugt größer 100 Dalton, besonders bevorzugt größer 1000 Dalton, ganz besonders bevorzugt größer 200 Dalton haben. Vorteilhaft sollten die niedermolekularen Substanzen weniger als drei Hydroxylgruppen an einem Kohlenstoffatom-enthaltenden Ring haben. Weiterhin sollten keine freie(n) Säure- oder Lacton-Gruppe(n) sowie keine
5 Phosphatgruppe und nicht mehr als eine Aminogruppe im Molekül enthalten sein. Auch Basen wie Adenosin im Molekül sind weniger bevorzugt.

In einer vorteilhaften Ausführungsform der Substanzen handelt es sich bei der Substanz um eine proteinogene Substanz, um eine Antisense-RNA, eine inhibierende
10 oder eine interferierende RNA (RNAi).

Der Begriff "sense" bezieht sich auf den Strang einer doppelsträngigen DNA der homolog zu dem mRNA-Transkript ist. Der "Antisense"-Strang enthält eine invertierte Sequenz, die komplementär zu der des "Sense"-Strangs ist. Ein Antisense-Nukleinsäuremolekül umfasst z.B. eine Nukletidsequenz, die komplementär zu dem "Sense"-
15 Nukleinsäuremolekül ist, das ein Protein oder eine aktive RNA codiert, z.B. komplementär zu dem codierenden Strang eines doppelsträngigen cDNA-Moleküls oder komplementär zu einer mRNA-Sequenz. Folglich kann ein Antisense-Nukleinsäuremolekül Wasserstoffbrückenbindungen zu einem Sense-Nukleinsäuremolekül ausbilden. Das Antisense-Nukleinsäuremolekül kann komplementär zu jedem hier
20 gezeigten codierenden Strang sein, oder nur zu einem Teil davon. Der Begriff "codierende Region" bezieht sich auf die Region einer Nukleinsäuresequenz, deren Codone in Aminosäuren translatiert werden. Auch kann das Antisense-Nukleinsäuremolekül komplementär zu "nicht-codierenden Regionen" des codierenden Strangs der gezeigten Nukleinsäuremoleküle sein. Der Begriff "nicht-codierende Region" bezieht sich
25 auf 5'- und 3'-Sequenzen, die die codierende Region flankieren und die nicht in ein Polypeptid translatiert werden (z.B. auch bezeichnet als 5'- und 3'-nicht-translatierte Regionen). Das Nukleinsäuremolekül, das eine Antisensesequenz umfasst, kann auch weitere für die Expression und Stabilität des Moleküls wichtige Elemente umfassen,
30 z.B. Capping-Strukturen, poly A-tails etc.

Das Antisense-Nukleinsäuremolekül kann komplementär zu der gesamten codierenden Region einer mRNA sein, aber es kann auch ein Oligonukleotid sein, welches komplementär zu nur einem Teil der codierenden oder nicht-codierenden Region der mRNA
35 ist. Z.B. kann ein Antisense-Oligonukleotid komplementär zu der Region sein, die den Translationsstart der mRNA umfasst oder umgibt. Ein Antisense-Oligonukleotid kann vorteilhaft z.B. 10-, 15-, 20-, 25-, 30-, 35-, 40-, 45- oder 50-Nukleotide lang sein. Ein Antisense-Nukleinsäuremolekül kann durch chemische Synthese und enzymatische Ligation nach dem Fachmann bekannten Verfahren hergestellt werden. Eine Anti-
40 sense-Nukleinsäuremolekül kann chemisch synthetisiert werden unter Verwendung

von natürlich vorkommenden Nukleotiden oder auf verschiedene Arten modifizierten Nukleotiden, so dass die biologische Stabilität der Moleküle erhöht ist oder die physikalische Stabilität des Duplex, die sich zwischen der Antisense- und Sense-Nukleinsäure bildet, verstärkt ist, z.B. können Phosphorothioatderivate und Acridin-substituierte Nukleotide verwendet werden. Beispiele für modifizierte Nukleotide, die für die Herstellung von Antisense-Nukleinsäuren verwendet werden können, umfassen 5-Fluorouracil, 5-Bromouracil, 5-Chlorouracil, 5-Iodouracil, Hypoxanthine, Xanthine, 4-Acetylcytosine, 5-(Carboxyhydroxymethyl)uracil, 5-Carboxymethylaminomethyl-2-thiouridine, 5-Carboxymethylaminomethyluracil, Dihydrouracil, Beta-D-galactosyl-queosine, Inosine, N6-Isopentenyladenine, 1-Methylguanine, 1-Methylinosine, 2,2-Dimethylguanine, 2-Methyladenine, 2-Methylguanine, 3-Methylcytosine, 5-Methylcytosine, N6-Adenine, 7-Methylguanine, 5-Methylaminomethyluracil, 5-Methoxyaminomethyl-2-thiouracil, Beta-D-mannosylqueosine, 5'-Methoxycarboxymethyluracil, 5-Methoxyuracil, 2-Methylthio-N6-isopentenyladenine, Uracil-5-oxyacetic acid (v), Wybutoxosine, Pseudouracil, Queosine, 2-Thiocytosine, 5-Methyl-2-thiouracil, 2-Thiouracil, 4-Thiouracil, 5-Methyluracil, Uracil-5-oxyaceticacidmethylester, Uracil-5-oxyacetic acid (v), 5-Methyl-2-thiouracil, 3-(3-Amino-3-N-2-carboxypropyl)uracil, (acp3)w, und 2,6-Diaminopurine.

Alternativ können Antisense-Nukleinsäuremoleküle biologisch hergestellt werden unter Verwendung von Expressionsvektoren, in welche Polynukleotide kloniert wurden, deren Orientation gegenläufig ist (so dass RNA, transkribiert von dem inserierten Polynukleotid, in einer Antisenseorientierung zu einem Zielpolynukleotid wie es weiter oben beschrieben wurde, ist).

Das Antisense-Nukleinsäuremolekül kann auch ein "α-Anomeric" Nukleinsäuremolekül sein. Ein "α-Anomeric" Nukleinsäuremolekül formt spezifische Doppelstranghybride mit komplementären RNAs, in denen, im Gegensatz zu gewöhnlichen β-Einheiten, die Stränge parallel zueinander laufen. Das Antisense-Nukleinsäuremolekül kann 2-O-Methylribonukleotide oder chimäre RNA-DNA-Analoga umfassen.

Weiterhin kann das Antisense-Nukleinsäuremolekül ein Ribozym sein. Ribozyme sind katalytische RNA-Moleküle mit einer Ribonukleaseaktivität, die dazu fähig sind, einzelsträngige Nukleinsäuren, wie z.B. mRNA, zu denen sie eine komplementäre Region haben, zu schneiden. Ribozyme (z.B. Hammerheadribozyme) können dazu verwendet werden katalytisch oder nichtkatalytisch mRNA der hierin beschriebenen Sequenzen zu schneiden und somit die Translation der mRNA zu verhindern. Ein Ribozym, das zu einer der hierin genannten Nukleinsäuresequenzen spezifisch ist, kann aufgrund der hier gezeigten cDNA-Sequenzen konstruiert werden oder auf Basis von heterologen Sequenzen, die nach den hierin beschriebenen Methoden identifiziert

werden können. B. kann ein Derivat der Tetrahymena L-15 dsRNA hergestellt werden, indem die Nukletidsequenz der aktiven Region komplementär zu der Nukleotidsequenz ist, die in einer codierenden mRNA geschnitten wird. Alternativ kann auch eine der hierin beschriebenen codierenden oder nicht-codierenden Sequenzen oder einer mRNA davon verwendet werden, um eine katalytische RNA aus einem Pool von RNAs auszuwählen (s. z.B. Bartel, 1993, Science, 261, 1411). Alternativ kann die Expression auch inhibiert werden, indem Nukleotidsequenzen, die komplementär zu einer regulatorischen Region der hierin beschriebenen Nukleinsäuresequenzen ist (z.B. ein Promotor oder Enhancer) eine triple-helikale Struktur bildet, die die Transkription des folgenden Gens verhindert (z.B. Helene, 1991, Anticancer-Drug Des. 6, 596; Helene, 1992, Ann. NY Acad. Sci. 660, 27, oder Maher, 1992, Bioassays, 14, 807.

Die dsRNAi-Methode (= "double-stranded RNA interference") wurde vielfach in tierischen und pflanzlichen Organismen beschrieben (z.B. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). Auf die in den angegebenen Zitaten beschriebenen Verfahren und Methoden wird ausdrücklich Bezug genommen. Eine effiziente Gensuppression kann auch bei transienter Expression oder nach transienter Transformation beispielsweise infolge einer biolistischen Transformation gezeigt werden (Schweizer P et al. (2000) Plant J 2000 24: 895-903). dsRNAi-Verfahren beruhen auf dem Phänomen, dass durch gleichzeitiges Einbringen von komplementären Strang- und Gegenstrang eines Gentranskriptes eine hocheffiziente Unterdrückung der Expression des entsprechenden Gens bewirkt wird. Der bewirkte Phänotyp kommt dem einer entsprechenden knock-out Mutanten sehr ähnlich (Waterhouse PM et al. (1998) Proc Natl Acad Sci USA 95:13959-64).

Das dsRNAi-Verfahren kann vorteilhaft zur Verminderung der Expression der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, ihrer Derivate und Fragmente verwendet werden. Wie u.a. in WO 99/32619 beschrieben sind dsRNAi-Ansätze klassischen antisense-Ansätzen deutlich überlegen.

Ein weiterer Gegenstand der Erfindung bezieht sich daher auf doppelsträngige RNA-Moleküle (dsRNA-Moleküle), die bei Einführung in einen Organismus vorteilhaft einer Pflanze (oder eine davon abgeleitete Zelle, Gewebe, Organ oder Samen) die Ver-

- minderung der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, ihrer Derivate oder Fragmente oder der durch sie codierten Proteine bewirken. Das doppelsträngiges RNA-Molekül zur Verminderung der Expression eines Proteins, das durch die Sequenzen SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 codiert ist, ist dadurch gekennzeichnet, dass
- i) einer der beiden RNA Stränge im wesentlichen identisch ist zu zumindest einem Teil einer Nukleinsäuresequenz mit den Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, und
 - ii) der jeweils andere RNA Strang im wesentlichen identisch ist zu zumindest einem Teil des komplementären Stranges einer der unter (i) genannten Nukleinsäuresequenz.
- "Im wesentlichen identisch" meint, dass die dsRNA Sequenz auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu der Zielsequenz (SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51) aufweisen kann und dennoch eine effizient Verminderung der Expression bewirken. Bevorzugt beträgt die Homologie nach obiger Definition mindestens 75 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 % am meisten bevorzugt 100 % zwischen dem "sense"-Strang einer inhibitorischen dsRNA und einem Teilabschnitt einer Nukleinsäuresequenz mit den Sequenzen

SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9,
SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19,
SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29,
SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,
5 SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49
oder SEQ ID NO: 51(bzw. zwischen dem "antisense"-Strang dem komplementären
Strang einer Nukleinsäuresequenz der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3,
SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13,
SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23,
10 SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33,
SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43,
SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51). Die Länge des
Teilabschnittes beträgt mindestens 10 Basen, bevorzugt mindestens 25 Basen,
besonders bevorzugt mindestens 50 Basen, ganz besonders bevorzugt mindestens
15 100 Basen, am meisten bevorzugt mindestens 200 Basen oder mindestens 300 Basen.
Alternativ, kann eine "im wesentlichen identische" dsRNA auch als Nukleinsäure-
sequenz definiert werden, die befähigt ist, mit einem Teil eines Gentranskriptes
der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7,
SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17,
20 SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27,
SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37,
SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
SEQ ID NO: 49 oder SEQ ID NO: 51 zu hybridisieren (z.B. in 400 mM NaCl, 40 mM
PIPES pH 6,4, 1 mM EDTA bei 50°C oder 70°C für 12 bis 16 h).

25 Die dsRNA kann aus einem oder mehr Strängen polymerisierter Ribonukleotide
bestehen. Es können ferner Modifikationen sowohl des Zucker-Phosphat-Gerüsts als
auch der Nukleoside vorliegen. Beispielsweise können die Phosphodiesterbindungen
der natürlichen RNA dahingehend modifiziert sein, dass sie zumindest ein Stickstoff-
30 oder Schwefel-Heteroatom umfassen. Basen können dahingehend modifiziert werden,
dass die Aktivität beispielsweise von Adenosindeaminase eingeschränkt wird. Solche
und weitere Modifikationen sind weiter unten bei den Verfahren zur Stabilisierung von
antisense-RNA beschrieben.

35 Die dsRNA kann enzymatisch oder ganz oder teilweise chemisch-synthetisch her-
gestellt werden.

Die doppelsträngige Struktur kann ausgehend von einem
einzelnen, selbstkomplementären Strang oder ausgehend von zwei komplementären
40 Strängen gebildet werden. Bei einem einzelnen, selbstkomplementären Strang, können

"sense"- und "antisense"-Sequenz durch eine verbindende Sequenz ("Linker") verknüpft sein und beispielsweise eine Haarnadelstruktur ausbilden. Bevorzugt kann die verbindende Sequenz ein Intron sein, das nach Synthese der dsRNA herausgespleißt wird. Die Nukleinsäuresequenz codierend für eine dsRNA kann weitere Elemente beinhalten, wie beispielsweise Transkriptionsterminationssignale oder Polyadenylierungssignale. Sollen die zwei Stränge der dsRNA in einer Zelle oder einem Organismus vorteilhaft einer Pflanze zusammengebracht werden, so kann dies auf verschiedene Art geschehen:

- 10 a) Transformation der Zelle oder des Organismus vorteilhaft einer Pflanze mit einem Vektor, der beide Expressionskassetten umfasst,
- b) Kotransformation (= Co-Transformation) der Zelle oder des Organismus vorteilhaft einer Pflanze mit zwei Vektoren, wobei der eine die Expressionskassetten mit dem "sense"-Strang, der andere die Expressionskassetten mit dem "antisense"-Strang umfasst.
- 15 c) Kreuzung von zwei Organismen vorteilhaft von Pflanzen, die mit jeweils einem Vektor transformiert wurden, wobei der eine die Expressionskassetten mit dem "sense"-Strang, der andere die Expressionskassetten mit dem "antisense"-Strang umfasst.
- 20

Die Bildung der RNA Duplex kann entweder außerhalb der Zelle oder innerhalb derselben initiiert werden. Wie in WO 99/53050 kann die dsRNA auch eine Haarnadelstruktur umfassen, indem "sense"- und "antisense"-Strang durch einen "Linker" (beispielsweise ein Intron) verbunden werden. Die selbstkomplementären dsRNA-Strukturen sind bevorzugt, da sie lediglich die Expression eines Konstruktes erfordern und die komplementären Stränge stets in einem äquimolaren Verhältnis umfassen.

30 Die Expressionskassetten codierend für den "antisense"- oder "sense"-Strang einer dsRNA oder für den selbstkomplementären-Strang der dsRNA, werden bevorzugt in einen Vektor inseriert und mit den unten beschriebenen Verfahren stabil (beispielsweise unter Verwendung von Selektionsmarkern) in das Genom einer Pflanze inseriert, um eine dauerhafte Expression der dsRNA zu gewährleisten.

35 Die dsRNA kann unter Verwendung einer Menge eingeführt werden, die zumindest ein Kopie pro Zelle ermöglicht. Höhere Mengen (z.B. mindestens 5, 10, 100, 500 oder 1000 Kopien pro Zelle) können ggf. eine effizienter Verminderung bewirken.

- Wie bereits beschrieben, ist eine 100%ige Sequenzidentität zwischen dsRNA und einem Gentranskript der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 nicht zwingend erforderlich, um eine effiziente Verminderung der Expression der genannten Sequenzen zu bewirken. Demzufolge besteht der Vorteil, dass das Verfahren tolerant ist gegenüber Sequenzabweichungen, wie sie infolge genetischer Mutationen, Polymorphismen oder evolutionärer Divergenzen vorliegen können. So ist es beispielsweise möglich mit der dsRNA, die ausgehend von den Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 des einen Organismus generiert wurde, die Expression der Sequenzen in einem anderen Organismus zu unterdrücken. Die hohe Sequenzhomologie zwischen den Sequenzen aus verschiedenen Organismen lässt auf einen hohen Konservierungsgrad dieser Proteine innerhalb von beispielsweise Pflanzen schließen, so dass die Expression einer dsRNA abgeleitet von einer der offenbarten Sequenzen gemäß SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 auch einen vorteilhaften Effekt in anderen Pflanzenarten haben dürfte.
- Die dsRNA kann entweder in vivo oder in vitro synthetisiert werden. Dazu kann eine DNA-Sequenz codierend für eine dsRNA in eine Expressionskassette unter Kontrolle mindestens eines genetischen Kontrollelementes (wie beispielsweise Promotor, Enhancer, Silencer, Splice-Donor oder -Akzeptor, Polyadenylierungssignal) gebracht werden. Entsprechend vorteilhafte Konstruktionen sind weiter unten beschrieben. Eine Polyadenylierung ist nicht erforderlich, ebenso müssen keine Elemente zur Initiierung einer Translation vorhanden sein.

- Eine dsRNA kann chemisch oder enzymatisch synthetisiert werden. Dazu können zelluläre RNA Polymerasen oder Bakteriophagen RNA Polymerasen (wie z.B. T3-, T7- oder SP6 RNA-Polymerase) verwendet werden. Entsprechende Verfahren zu in vitro

Expression von [REDACTED] sind beschrieben (WO 97/32016; US 5, [REDACTED] 874; US 5,698,425, US 5,712,135, US 5,789,214, US 5,804,693). Eine chemisch oder enzymatisch in vitro syntetisierte dsRNA kann vor der Einführung in eine Zelle, Gewebe oder Organismus aus dem Reaktionsgemisch beispielsweise durch Extraktion, Präzipitation, Elektrophorese, Chromatographie oder Kombinationen dieser Verfahren ganz oder teilweise aufgereinigt werden. Die dsRNA kann unmittelbar in die Zelle eingeführt werden oder aber auch extrazellulär (z.B. in den interstitialen Raum) appliziert werden.

Unter "Antikörpern" sind beispielsweise polyklonale, monoklonale, humane oder humanisierte oder rekombinante Antikörper oder Fragmente davon, single chain Antikörper oder auch synthetische Antikörper zu verstehen. Unter erfindungsgemäßen Antikörpern oder deren Fragmente sind prinzipiell alle Immunglobulinklassen wie IgM, IgG, IgD, IgE, IgA oder ihre Subklassen wie die Subklassen des IgG oder deren Mischungen zu verstehen. Bevorzugt sind IgG und seine Subklassen wie beispielsweise IgG1, IgG2, IgG2a, IgG2b, IgG3 oder IgGM. Besonders bevorzugt sind die IgG Subtypen IgG1 oder IgG2b. Als Fragmente seien alle verkürzten oder veränderten Antikörperfragmente mit einer oder zwei dem Antigen komplementären Bindungsstellen, wie Antikörperteile mit einer den Antikörper entsprechenden von leichter und schwerer Kette gebildeten Bindungsstelle wie Fv-, Fab- oder F(ab')₂-Fragmente oder Einzelstrangfragmente, genannt. Bevorzugt sind verkürzte Doppelstrangfragmente wie Fv-, Fab- oder F(ab')₂. Diese Fragmente können beispielsweise auf enzymatischem Wege durch Abspaltung des Fc-Teils der Antikörper mit Enzymen wie Papain oder Pepsin, durch chemische Oxidation oder durch gentechnische Manipulation der Antikörpergene erhalten werden. Auch genmanipulierte nichtverkürzte Fragmente können vorteilhaft verwendet werden. Die Antikörper oder Fragmente können allein oder in Mischungen verwendet werden. Antikörper können auch Teil eines Fusionsproteins sein.

Die identifizierten Substanzen können chemisch synthetisierte oder mikrobiologisch produzierte Stoffe sein und z.B. in Zellextrakten von z.B. Pflanzen, Tieren oder Mikroorganismen auftreten. Weiterhin können die genannten Stoffe zwar im Stand der Technik bekannt sein, aber bisher nicht bekannt sein als Herbizid. Das Reaktionsgemisch kann ein zellfreier Extrakt sein oder eine Zelle oder Zellkultur umfassen. Geeignete Methoden sind dem Fachmann bekannt und werden z.B. allgemein beschrieben in Alberts, Molecular Biology the cell, 3rd Edition (1994), z.B. Kapitel 17. Die genannten Stoffe können z.B. zu dem Reaktionsgemisch oder dem Kulturmedium zugegeben werden oder den Zellen injiziert werden oder auf eine Pflanze gesprüht werden.

- Wenn eine Probe eine nach der erfindungsgemäßen Methode aktive Substanz beinhaltet, identifiziert wurde, dann ist es entweder möglich, den Stoff direkt von der ursprünglichen Probe zu isolieren oder man kann die Probe in verschiedene Gruppen teilen, z.B. wenn sie aus einer Vielzahl von verschiedenen Komponenten besteht, um so die Zahl der verschiedenen Substanzen pro Probe zu reduzieren und dann das erfindungsgemäße Verfahren mit einer solchen "Unterprobe" der ursprünglichen Probe zu wiederholen. Abhängig von der Komplexität der Probe können die oben beschriebenen Schritte mehrmals wiederholt werden, vorzugsweise bis die gemäß der erfindungsgemäßen Methode identifizierte Probe nur noch eine geringe Anzahl von Substanzen oder nur noch eine Substanz umfasst. Vorzugsweise wird der gemäß der erfindungsgemäßen Methode identifizierte Stoff oder Derivate davon weiter formuliert, so, dass er für die Anwendung in der Pflanzenzüchtung oder Pflanzenzell- oder Gewebekultur geeignet ist.
- Die Stoffe, die gemäß dem erfindungsgemäßen Verfahren getestet und identifiziert wurden, können beispielsweise sein: Expressionsbibliotheken, z.B. cDNA-Expressionsbibliotheken, Peptide, Proteine, Nukleinsäuren, Antikörper, kleine organische Stoffe, Hormone, PNAs oder ähnliches (Milner, Nature Medicine 1 (1995), 879-880; Hupp, Cell. 83 (1995), 237-245; Gibbs, Cell. 79 (1994), 193-198 und darin zitierte Referenzen).
- Diese Stoffe könne auch funktionelle Derivate oder Analogon der bekannten Inhibitoren oder Aktivatoren sein. Verfahren zur Herstellung von chemischen Derivaten oder Analogon sind dem Fachmann bekannt. Die genannten Derivate und Analogon können gemäß Verfahren nach dem Stand der Technik getestet werden. Weiterhin kann computergestütztes Design oder Peptidomimetics zur Herstellung geeigneter Derivate und Analogon verwendet werden. Die Zelle oder das Gewebe, die/das für das erfindungsgemäße Verfahren verwendet werden kann, ist vorzugsweise eine erfindungsgemäße Wirtszelle, Pflanzenzelle oder ein Pflanzengewebe, wie in den oben genannten Ausführungsformen beschrieben.
- Unter Derivate(n) (der Plural und der Singular seien für diese Anmeldung und deren Definitionen äquivalent) der in den erfindungsgemäßen Verfahren verwendeten Nukleinsäuren sind beispielsweise funktionelle Homologe der von SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 codierten Proteine oder deren biologischer Aktivität, das heißt Proteine, die dieselben biologischen Reaktionen wie die von SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,

SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 codierten Proteine ausführen, zu verstehen. Diese Derivate bzw. Gene sind ebenfalls als herbizide Targets geeignet.

Die hierin erfindungsgemäß beschriebenen Sequenzen codieren für Homologe zu den in den Beispielen beschriebenen Proteinen und haben vorzugsweise die für die Homologe angegebenen Aktivitäten.

10

SEQ ID NO: 1 codiert für ein Protein, das zum Translation Realising Factor RF-2 Ähnlichkeiten aufweist. Die Proteinsequenz ist in SEQ ID NO: 2 wiedergegeben.

SEQ ID NO: 3 codiert für ein Kobalamin-Synthese Protein, dessen Proteinsequenz

SEQ ID NO: 4 zu entnehmen ist. SEQ ID NO: 5 codiert für eine Arginyl-tRNA-

15

Synthetase, die Proteinsequenz ist in SEQ ID NO: 6 dargestellt. SEQ ID NO: 7 codiert für ein putatives Protein mit Ähnlichkeit zu einer RNA-Helicase aus *Mus musculus*, dessen Proteinsequenz in SEQ ID NO: 8 wiedergegeben ist. SEQ ID NO: 9 codiert für ein putatives Protein mit Ähnlichkeit zu dem die AP2-Domäne enthaltenden Protein RAP 2.4 aus *Arabidopsis thaliana*, dessen Proteinsequenz SEQ ID NO: 10 zu entnehmen

20

ist. SEQ ID NO: 11 codiert für ein Protein, das Homologien zu verschiedenen Pseudouridylat-Synthasen aufweist. Die Proteinsequenz ist SEQ ID NO: 12 zu entnehmen. Von SEQ ID NO: 13 wird ein Protein codiert, das Ähnlichkeiten mit einer putativen Adenylatkinase hat. SEQ ID NO: 14 gibt die Proteinsequenz wieder. Für ein Protein, mit der in SEQ ID NO: 16 gezeigten Sequenz, codiert die Sequenz

25

SEQ ID NO: 15. Dieses durch SEQ ID NO: 15 codierte hypothetische Protein hat Ähnlichkeit zu dem pol-Polypeptid des "Equine Infectious Anemia Virus".

SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 35, SEQ ID NO: 43 und SEQ ID NO: 51 codieren für unbekannte Proteine. Den Sequenzen SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 36, SEQ ID NO: 44 und

30

SEQ ID NO: 52 sind die jeweiligen Proteinsequenzen zu entnehmen.

SEQ ID NO: 23 codiert für ein Preprotein Translocase secA Precursor-Protein. Es handelt sich also um ein chloroplastidäres SecA-Protein, das am Proteintransport über die Thylakoidmembran beteiligt ist. Die Proteinsequenz ist SEQ ID NO: 24 zu entnehmen.

35

Von SEQ ID NO: 25 wird ein Protein mit signifikanter Homologie zum DCL-Protein aus Tomate (PIR: S71749) codiert. Dieses Protein besitzt eine sogenannte HMG-Signatur, die in high mobility group Proteinen vorkommt und die an DNA binden kann. Die

40

Proteinsequenz wird in SEQ ID NO: 26 wiedergegeben.

Durch SEQ ID NO: 29 wird eine plastidäre Gluthathion-Reduktase codiert, deren Proteinsequenz in SEQ ID NO: 30 dargestellt wird. SEQ ID NO: 31 codiert für ein Protein, das ein Homolog des Transkriptionsfaktors Sigma ist, d.h. es ist ein pflanzliches Homolog zur Sigma-Untereinheit der bakteriellen RNA-Polymerase. Die
5 entsprechende Proteinsequenz ist SEQ ID NO: 32 zu entnehmen.

Von SEQ ID NO: 33 wird ein Calmodulin-ähnliches Protein codiert, deren Sequenz in
10 SEQ ID NO: 34 dargestellt wird.

SEQ ID NO: 37 codiert für ein Protein mit großer Ähnlichkeit zu INT6, einem
brustkrebsassoziierten Protein, mit mit Ähnlichkeit zu einem „initiationfaktor 3“ Protein.
SEQ ID NO: 38 gibt die Proteinsequenz wieder.

15 Von SEQ ID NO: 39 wird ein Protein codiert, das große Ähnlichkeit zur DNA Helikase
YGL150c aus *Saccharomyces* hat. SEQ ID NO: 40 stellt die entsprechende Proteinsequenz dar.

Ein Protein mit Ähnlichkeit zu einem RNA-bindenden Protein wird durch
20 SEQ ID NO: 41 codiert. Die Proteinsequenz wird in SEQ ID NO: 42 wiedergegeben.

SEQ ID NO: 45 codiert für einen putativen Hitzeschock-Transkriptionsfaktor, dessen
Proteinsequenz SEQ ID NO: 46 zu entnehmen ist.

25 Von SEQ ID NO: 47 wird ein putatives, chloroplastidäres an das DNA-Nukleoid
bindende Protein codiert. SEQ ID NO: 48 gibt die entsprechende Proteinsequenz
wieder.

30 SEQ ID NO: 49 codiert für ein Protein, das Ähnlichkeit zu einer putativen Met2-Typ
Cytosin DNA-Methyltransferase hat. Diese Methyltransferase hat starke Ähnlichkeiten
zu einer DNA(cytosine-5-)-Methyltransferase aus *Arabidopsis thaliana*. Die Proteinsequenz ist in SEQ ID NO: 50 dargestellt.

35 Unter Derivaten werden auch solche Peptide verstanden, die eine Homologie zu den
Polypeptiden mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8,
SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18,
SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28,
SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38,
40 SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48,

- SEQ ID NO: 50 oder SEQ ID NO: 52 gezeigten Sequenzen vorzuziehen, mindestens 20 %, vorzugsweise 30 %, 40 % oder 50 %, mehr bevorzugt 60 %, 70 % oder 80 %, noch mehr bevorzugt 90 %, mehr bevorzugt 91 %, 92 %, 93 %, 94 % oder 95 %, am meisten bevorzugt 96 %, 97 %, 98 % oder 99 % oder mehr haben, und die eine äquivalente biologische Aktivität in anderen Organismen besitzen und somit als funktionelle Homologe anzusehen sind. Diese funktionelle Homologie oder Äquivalenz lässt sich z.B. durch die mögliche Komplementation von Mutanten in diesen Funktionen demonstrieren.
- 5 Die oben genannten Nukleinsäuresequenz(en) oder Fragmente davon können vorteilhaft zur Isolierung weiterer Sequenzen wie z.B. genomische, cDNA oder sonstige Sequenzen, die als Herbizidtarget geeignet sind, über Homologiescreening verwendet werden.
- 10 Die genannten Derivate lassen sich beispielsweise aus anderen Organismen, insbesondere eukaryontischen Organismen wie monokotylen oder dikotylen Pflanzen, wie speziell Algen, Moosen, Dinoflagellaten, Nutzpflanzen wie Monokotyle wie Mais, Weizen, Hafer, Roggen, Gerste oder Hirse oder Dikotyle wie Kartoffel, Tabak, Salat, Tomate, Karotte um nur einige zu nennen oder Pilze, isolieren.
- 15 Weiterhin sind unter Derivaten bzw. funktionellen Derivaten der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 genannten Sequenzen beispielsweise Allelvarianten zu verstehen, die mindestens 60 % Homologie auf der abgeleiteten Aminosäureebene, vorteilhaft mindestens 70 % Homologie, bevorzugt mindestens 80 % Homologie, besonders bevorzugt mindestens 85 %, 90 %, 91 %, 92 %, 93 %, 94 % oder 95 % Homologie, ganz besonders bevorzugt 96 %, 97 %, 98 % oder 99 % Homologie aufweisen. Die Homologie wurde über den gesamten Aminosäurebereich berechnet. Es wurde das Programm PileUp, BESTFIT, GAP, TRANSLATE bzw. BACKTRANSLATE (= Bestandteil des Programmpaketes UWGCG, Wisconsin Package, Version 10.0-UNIX, January 1999, Genetics Computer Group, Inc., Deverux et al., Nucleic. Acid Res., 12, 1984: 387-395) verwendet (J. Mol. Evolution., 25, 351-360, 1987; Higgins et al., CABIOS, 5 1989: 151-153). Für Nukleinsäuren wurden folgende Einstellungen verwendet: Gap Weight: 50, Length Weight: 3. Für Proteine waren die Einstellungen: Gap Weight: 8, Length Weight: 2. Die von den genannten Nukleinsäuren abgeleitete Aminosäuresequenzen sind Sequenz
- 20
- 30
- 35
- 40
- SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10,

SEQ ID NO: 12, ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50
5 oder SEQ ID NO: 52 zu entnehmen. Unter Homologie ist Identität zu verstehen, das heißt die Aminosäuresequenzen sind zu mindestens 40, 50, 60 oder 70 %, mehr bevorzugt 80 %, 85 % oder 90 %, noch mehr bevorzugt 91 %, 92 %, 93 %, 94 % oder 95 %, am meisten bevorzugt 96 %, 97 %, 98 % oder 99 % oder mehr identisch. Die erfindungsgemäßen Sequenzen sind auf Nukleinsäureebene mindestens 45 oder 55 %
10 homolog, bevorzugt mindestens 60 oder 65 %, besonders bevorzugt 75 % oder 80 %, ganz besonders bevorzugt mindestens 85 % oder 90 %, noch mehr bevorzugt 95 %, 96 %, 97 %, 98 % oder 99 % oder mehr homolog.

Weiterhin umfasst der Begriff Derivate sowie der Begriff "Fragmente" auch Teilbereiche
15 oder Fragmente der aufgeführten Sequenzen oder deren homologen Sequenzen von mindestens 50 Aminosäuren, vorteilhaft von mindestens 40 Aminosäuren, bevorzugt von mindestens 30 Aminosäuren, besonders bevorzugt von mindestens 20 Aminosäuren, ganz besonders bevorzugt von mindestens 10 Aminosäuren, die es ermöglichen, selektiv interagierende Substanzen zu identifizieren. Der Begriff "Fragment",
20 "Sequenzfragment" oder "Teilsequenz" bedeutet eine verkürzte Sequenz der Originalsequenz. Die verkürzte Sequenz (Nukleinsäure oder Protein) kann unterschiedliche Längen haben, die minimale Sequenzlänge ist eine Sequenzlänge, die wenigstens eine vergleichbare Funktion, z.B. Bindungseigenschaften, oder Aktivität der Originalsequenz hat. Entsprechende Verfahren sind z.B. wie oben beschrieben SELDI, FCS
25 oder Biocore und sind dem Fachmann bekannt.

Ebenfalls umfasst sind somit Nukleinsäuren, die ein Fragment oder ein Epitope eines Polypeptides codieren, das spezifisch an einem Antikörper bindet, der spezifisch an eine, als erfindungsgemäß beschriebenes Polypeptid bindet, insbesondere das
30 von einer der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
35 SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz codiert wird. Fragment oder Epitope eines Polypeptides, die mit einem solchen Antikörper spezifisch interagieren, weisen eine signifikante Homologie in der räumlichen Struktur zu den hierin beschriebenen Polypeptiden, zumindest in Teilbereichen, auf. Vorzugsweise besitzen sie ebenfalls eine hohe Homologie auf Aminosäureebene zu den genannten Sequenzen,
40 vorzugsweise 20 %, mehr bevorzugt sind 40 %, mehr bevorzugt 60 %, noch mehr

80 %, am meisten bevorzugt sind 90 % oder mehr. Die räumliche Struktur eines Polypeptides ist jedoch im wesentlichen mitverantwortlich für die Interaktionen des Polypeptides mit anderen Verbindungen sowie ggf. für seine enzymatische Aktivität. Folglich können in den erfindungsgemäßen Verfahren oder Fragmenten eingesetzt werden, deren Sequenz nur eine geringe Homologie zu den beschriebenen Polypeptiden aufweist, deren räumliche Struktur jedoch eine hohe Homologie zu den beschriebenen Polypeptiden aufweist, also solche, die Epitope der hierin beschriebenen Sequenzen enthalten, um Interaktionspartner zu finden, die dann die hierin beschriebenen Polypeptide inhibieren oder inaktivieren. Fragmente, die Epitope der erfindungsgemäßen Polypeptide umfassen, können auch verwendet werden, um die Interaktionspartner der erfindungsgemäßen Polypeptide zu "besetzen", d.h. ihre Interaktion mit den erfindungsgemäßen Polypeptiden zu verhindern. Hierzu ist es dann vorteilhaft, wenn die Fragmente eine höhere Affinität zu einem Bindungspartner haben als das natürlich vorkommende Polypeptid. Ebenfalls umfasst sind Fragmente, die von erfindungsgemäßen Nukleinsäuren codiert werden und eine der oben genannten biologischen Aktivitäten umfassen.

Allelvarianten umfassen insbesondere funktionelle Varianten, die durch Deletion, Insertion oder Substitution von Nukleotiden aus der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz erhältlich sind, wobei die biologische, z.B. enzymatische Aktivität oder Bindungseigenschaften der abgeleiteten synthetisierten Proteine erhalten bleibt.

Solche DNA-Sequenzen lassen sich mit Hilfe der erfindungsgemäßen Nukleinsäuresequenzen, z.B. ausgehend von den in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 beschriebenen DNA-Sequenzen oder Teilen dieser Sequenzen, beispielsweise mit üblichen Hybridisierungsverfahren oder der PCR-Technik aus anderen Eukaryonten wie beispielsweise Mikroorganismen wie Hefen, Pilzen, Ciliaten, Pflanzen wie Algen, Moosen oder sonstigen Pflanzen, isolieren. Diese DNA-Sequenzen hybridisieren unter Standardbedingungen mit den genannten Sequenzen. Zur Hybridisierung werden vorteilhaft kurze Oligonukleotide, beispielsweise der konservierten oder sonstigen Bereiche, die über Vergleiche mit

anderen verwandten Genen in dem Fachmann bekannter Weise ermittelt werden können, verwendet. Es können aber auch längere Fragmente der erfindungsgemäßen Nukleinsäuren oder die vollständigen Sequenzen für die Hybridisierung verwendet werden. Je nach der verwendeten Nukleinsäure: Oligonukleotid, längeres Fragment oder vollständige Sequenz oder je nachdem welche Nukleinsäureart DNA oder RNA für die Hybridisierung verwendet werden, variieren diese Standardbedingungen. So liegen beispielsweise die Schmelztemperaturen für DNA:DNA-Hybride ca 10°C niedriger als die von DNA:RNA-Hybriden gleicher Länge.

- 10 Unter Standardbedingungen sind beispielsweise je nach Nukleinsäure Temperaturen zwischen 42 und 58°C in einer wässrigen Pufferlösung mit einer Konzentration zwischen 0,1 bis 5 x SSC (1 X SSC = 0,15 M NaCl, 15 mM Natriumcitrat, pH 7,2) oder zusätzlich in Gegenwart von 50 % Formamid wie beispielsweise 42°C in 5 x SSC, 50 % Formamid zu verstehen. Vorteilhafterweise liegen die Hybridisierungs-
- 15 bedingungen für DNA:DNA-Hybride bei 0,1 x SSC und Temperaturen zwischen etwa 20°C bis 45°C, bevorzugt zwischen etwa 30°C bis 45°C. Für DNA:RNA-Hybride liegen die Hybridisierungsbedingungen vorteilhaft bei 0,1 x SSC und Temperaturen zwischen etwa 30°C bis 55°C, bevorzugt zwischen etwa 45°C bis 55°C. Diese angegebenen Temperaturen für die Hybridisierung sind beispielhaft kalkulierte Schmelztemperatur-
- 20 werte für eine Nukleinsäure mit einer Länge von ca. 100 Nukleotiden und einem G + C-Gehalt von 50 % in Abwesenheit von Formamid. Die experimentellen Bedingungen für die DNA-Hybridisierung sind in einschlägigen Lehrbüchern der Genetik wie beispielsweise Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989, beschrieben und lassen sich nach dem Fachmann bekannten Formeln beispielsweise
- 25 abhängig von der Länge der Nukleinsäuren, der Art der Hybride oder dem G + C-Gehalt berechnen. Weitere Informationen zur Hybridisierung kann der Fachmann folgenden Lehrbüchern entnehmen: Ausubel et al. (eds), 1985, Current Protocols in Molecular Biology, John Wiley & Sons, New York; Hames and Higgins (eds), 1985, Nucleic Acids Hybridization: A Practical Approach, IRL Press at Oxford University
- 30 Press, Oxford; Brown (ed), 1991, Essential Molecular Biology: A Practical Approach, IRL Press at Oxford University Press, Oxford.

- Weiterhin sind unter Derivaten Homologe der Sequenz SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13,
- 35 SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 beispielsweise eukaryontische Homologe, verkürzte Sequenzen, Einzelstrang-DNA der codierenden

und nichtcodierende DNA-Sequenz oder RNA der codierenden und nichtcodierenden DNA-Sequenz zu verstehen.

Außerdem sind unter Homologen der Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 Derivate wie beispielsweise Varianten aus anderen Organismen beispielsweise anderen Pflanzen zu verstehen. Diese Varianten können durch ein oder mehrere Nukleotidaustausche, durch Insertion(en) und/oder Deletion(en) verändert sein, ohne dass aber die Funktionalität bzw. die biologische Aktivität der Varianten beeinträchtigt wird. Vorzugsweise haben sie eine Homologie von mindestens 20 %, vorteilhaft, 30 %, 40 %, 50 % oder 60 %, bevorzugt 70 %, 80 % oder 90 %, besonders bevorzugt von 95 % und eine äquivalente biologische Aktivität.

Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 und deren Fragmente und Derivate sind deshalb vorteilhaft dafür geeignet um weitere essentielle, neue Gene aus anderen Organismen bevorzugt Pflanzen zu isolieren.

Die erfindungsgemäßen Nukleinsäuresequenzen, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 und deren durch sie codierte Genprodukte werden im erfindungsgemäßen Verfahren verwendet. Sie können synthetisch hergestellt oder natürlich gewonnen sein oder eine Mischung aus synthetischen und natürlichen DNA-Bestandteilen enthalten, sowie aus verschiedenen heterologen Genabschnitten verschiedener Organismen bestehen. Im allgemeinen werden synthetische Nukleotid-Sequenzen mit Codons erzeugt, die von den entsprechenden Wirtsorganismen beispielsweise Pflanzen bevorzugt werden. Dies führt in der Regel zu einer optimalen Expression der heterologen Gene. Diese von

Pflanzen bevorzugt. Codons können aus Codons mit der hohen Proteinhäufigkeit bestimmt werden, die in den meisten interessanten Pflanzenspezies exprimiert werden. Ein Beispiel für *Corynebacterium glutamicum* ist gegeben in: Wada et al. (1992) Nucleic Acids Res. 20:2111-2118). Die Durchführung solcher Experimente sind mit Hilfe von Standardmethoden durchführbar und sind dem Fachmann auf dem Gebiet bekannt.

Funktionell äquivalente Sequenzen, die für die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren codieren, sind solche Derivate der erfindungsgemäßen Sequenzen, welche trotz abweichender Nukleotidsequenz noch die gewünschten Funktionen, das heißt die biologische Aktivität der Proteine besitzen. Funktionelle Äquivalente umfassen somit natürlich vorkommende Varianten der hierin beschriebenen Sequenzen sowie künstliche, z.B. durch chemische Synthese erhaltene, insbesondere an den Codon-Gebrauch einer Pflanze angepasste, künstliche Nukleotid-Sequenzen.

Außerdem sind artifizielle DNA-Sequenzen geeignet, solange sie, wie oben beschrieben, zu Produkten führen, die die oben genannten Aktivitäten oder die gewünschte Eigenschaft, beispielsweise die Bindung an einen Rezeptor oder die enzymatische Aktivität vermitteln. Solche artifiziellen DNA-Sequenzen können beispielsweise durch Rückübersetzung mittels Molecular Modelling konstruierter Proteine oder durch in vitro Selektion ermittelt werden. Mögliche Techniken zur in vitro-Evolution von DNA zur Veränderung bzw. Verbesserung der DNA-Sequenzen sind beschrieben bei Patten, P.A. et al., Current Opinion in Biotechnology 8, 724-733(1997) oder bei Moore, J.C. et al., Journal of Molecular Biology 272, 336-347(1997). Besonders geeignet sind codierende DNA-Sequenzen, die durch Rückübersetzung einer Polypeptidsequenz gemäß der für die Wirtspflanze spezifischen Codon-Nutzung erhalten werden. Die spezifische Codon-Nutzung kann ein mit pflanzen genetischen Methoden vertrauter Fachmann durch Computerauswertungen anderer, bekannter Gene der zu transformierenden Pflanze leicht ermitteln.

Für das erfindungsgemäße Verfahren sind vorteilhaft Aminosäuresequenzen zu verstehen, die eine in den Sequenzen SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellte Aminosäuresequenz oder eine daraus durch Substitution, Inversion, Insertion oder Deletion von einem oder mehreren Aminosäureresten erhältliche Sequenz enthalten, wobei die biologische

Aktivität des in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, 5 SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Proteins erhalten bleibt bzw. nicht wesentlich reduziert wird. Unter nicht wesentlich reduziert sind alle Proteine zu verstehen, die noch mindestens 10 %, bevorzugt 20 %, besonders bevorzugt 30 %, 50 %, 70 %, 90 % oder mehr, der biologischen Aktivität des Ausgangsproteins aufweisen. Dabei können beispielsweise bestimmte Aminosäuren durch solche mit 10 ähnlichen physikochemischen Eigenschaften (Raumerfüllung, Basizität, Hydrophobizität etc.) ersetzt werden. Beispielsweise werden Argininreste gegen Lysinreste, Valinreste gegen Isoleucinreste oder Asparaginsäurereste gegen Glutaminsäurereste ausgetauscht. Es können aber auch ein oder mehrere Aminosäuren in ihrer Reihen- 15 folge vertauscht, hinzugefügt oder entfernt werden, oder es können mehrere dieser Maßnahmen miteinander kombiniert werden.

Unter Derivaten sind auch funktionelle Äquivalente zu verstehen, die insbesondere auch natürliche oder künstliche Mutationen der verwendeten Nukleinsäuresequenzen. 20 SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 25 oder SEQ ID NO: 51 beinhalten, welche weiterhin die gewünschte Funktion, das heißt deren biologische Aktivität nicht wesentlich reduziert ist, zeigen. Mutationen umfassen Substitutionen, Additionen, Deletionen, Vertauschungen oder Insertionen eines oder mehrerer Nukleotidreste. Somit werden beispielsweise auch solche Nukleotidsequenzen durch die vorliegende Erfindung mit umfasst, welche man durch Modi- 30 fikation der genannten Nukleotidsequenzen erhält. Ziel einer solchen Modifikation kann z.B. die weitere Eingrenzung der darin enthaltenen codierenden Sequenz oder z.B. auch die Einfügung weiterer Restriktionsenzym-Schnittstellen sein.

Funktionelle Äquivalente sind auch solche Varianten, deren Funktion, verglichen mit 35 dem Ausgangsgen bzw. Genfragment, abgeschwächt (= nicht wesentlich reduziert) oder verstärkt ist (= Enzymaktivität stärker als die Aktivität des Ausgangsenzym, das heißt Aktivität ist höher als 100 %, bevorzugt höher als 150 %, besonders bevorzugt höher als 180 %).

Die Nukleinsäuresequenz kann dabei vorteilhaft beispielsweise eine DNA- oder cDNA-Sequenz sein. Zur Insertion in eine erfindungsgemäßen Nukleinsäurekonstrukt (= Expressionskassette oder Nukleinsäurefragment) geeignete codierende Sequenzen sind beispielsweise solche, die für ein Protein mit den oben beschriebenen Sequenzen codieren und die dem Wirt die Fähigkeit zur Überproduktion des Proteins und damit seiner biologischen Funktion verleihen. Diese Sequenzen können homologen oder heterologen Ursprungs sein.

Daher ist ein weiterer erfindungsgemäßer Gegenstand ein Nukleinsäurekonstrukt enthaltend eine erfindungsgemäße Nukleinsäuresequenz, z.B. ausgewählt aus der Gruppe:

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz, oder
- b) einer Nukleinsäuresequenz, die sich aufgrund des degenerierten genetischen Codes aus den durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen ableiten lässt,
- c) Nukleinsäuresequenz, die ein Derivat oder ein Fragment der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Nukleinsäuresequenzen ist, und mindestens 60 % Homologie auf Nukleinsäureebene aufweisen; oder

- d) Nukleinsäuresequenz, die für Derivate oder Fragmente der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen codiert, die mindestens 50 % Homologie auf Aminosäureebene aufweisen;
- e) Nukleinsäuresequenz, die für ein Fragment oder ein Epitope eines Polypeptides codiert, das spezifisch an einem Antikörper bindet, wobei der Antikörper spezifisch an ein Polypeptid bindet, das der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz codiert wird;
- f) Nukleinsäuresequenz, die ein Fragment einer in a) dargestellten Nukleinsäure codiert und das eine "Translation Releasing Factor"-Aktivität, eine Kobalamin-Synthese-Aktivität, Arginyl-tRNA-Synthase-Aktivität, eine RNA-Helicase-Aktivität, eine GTP-Bindeprotein-Aktivität, eine Pseudouridylat-Synthase-Aktivität, eine Adenylatkinase-Aktivität, eine Preprotein-Translocase secA Precursorprotein-Aktivität, eine DCL-Protein-Aktivität, eine Arginin-tRNA-Ligase-Aktivität, eine plastidäre Gluthathion-Reduktase-Aktivität, eine Transkriptionsfaktor Sigma-Aktivität, eine Calmodulin-Aktivität, eine INT6-Aktivität, eine Helikase YGL150c-Aktivität, eine RNA-bindende Aktivität, eine Hitzeschock-Transkriptionsfaktor-Aktivität, eine chloroplastidäre DNA-Nukleoid bindende Aktivität oder eine Met2-Typ Cytosin DNA Methyltransferase-Aktivität hat; und/oder
- g) Nukleinsäuresequenz, die für Derivate der Polypeptide mit den SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36,

SEQ ID NO: 3, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44,
SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52
dargestellten Aminosäuresequenzen codiert, die mindestens 20 % Homologie
auf Aminosäureebene aufweist und eine äquivalente biologische Aktivität besitzt;

5

wobei die Nukleinsäuresequenz mit einem oder mehreren Regulationssignalen
verknüpft ist. Die vorhergenannten Begriffe haben die oben genannte Bedeutung.

Unter dem erfindungsgemäßen Nukleinsäurekonstrukt sind die erfindungsgemäßen
10 Nukleinsäuren, z.B. die in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5,
SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,
SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25,
SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35,
SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45,
15 SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 genannten Sequenzen, die
sich als Ergebnis des genetischen Codes und/oder deren funktionellen oder nicht
funktionellen Derivate zu verstehen, die mit einem oder mehreren Regulationssignalen
vorteilhafterweise zur Regulation, insbesondere Erhöhung, der Genexpression
funktionell verknüpft wurden und welche die Expression der codierenden Sequenz in
20 der Wirtszelle steuern. Diese regulatorischen Sequenzen sollen die gezielte Expressi-
on der Gene bzw. der Proteine ermöglichen. Dies kann beispielsweise je nach
Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert und/oder
überexprimiert wird, oder dass es konstitutiv exprimiert und/oder überexprimiert wird.
Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen
25 an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure
regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder anstelle dieser
Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen
Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden
sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene
30 erhöht wurde. Das erfindungsgemäße Nukleinsäurekonstrukt kann auch vorteilhaft nur
aus der natürlichen gentechnologisch veränderten Regulationsregion am 5'- und/oder
3'-Ende bestehen. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das
heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz
oder dessen Derivate inseriert und der natürliche Promotor mit seiner Regulation
35 wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert,
dass keine Regulation mehr erfolgt und/oder die Genexpression gesteigert wird. Diese
veränderten Promotoren können in Form von Teilsequenzen (= Promotor mit Teilen
der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen
zur Steigerung der Aktivität gebracht werden. Das Genkonstrukt kann außerdem
40 vorteilhafterweise auch eine oder mehrere sogenannte "enhancer Sequenzen"

funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen können in einer oder mehreren Kopien in der Expressionskassette (= Genkonstrukt) enthalten sein.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird. In einer anderen vorteilhaften Ausführungsform kann die Expression jedoch auch gezielt reduziert oder blockiert werden.

Als Promotoren in der Expressionskassette sind grundsätzlich alle Promotoren geeignet, die die Expression von Fremdgenen in Organismen vorteilhaft in Pflanzen oder Pilzen steuern können. Vorzugsweise verwendet man insbesondere ein pflanzliche Promotoren oder Promotoren, die aus einem Pflanzenvirus entstammen. Vorteilhafte Regulationssequenzen für das erfindungsgemäße Verfahren sind beispielsweise in Promotoren wie *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacI^q*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, λ -*P_R*- oder im λ -*P_L*-Promotor enthalten, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Weitere vorteilhafte Regulationssequenzen sind beispielsweise in den gram-positiven Promotoren *amy* und *SPO2*, in den Hefe- oder Pilzpromotoren *ADC1*, *MF α* , *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH* oder in den Pflanzenpromotoren wie *CaMV/35S* [Franck et al., *Cell* 21(1980) 285-294], *SSU*, *OCS*, *lib4*, *STLS1*, *B33*, *nos* (= Nopalin Synthase Promotor) oder im Ubiquitin-Promotor enthalten. Die Expressionskassette kann auch einen chemisch induzierbaren Promotor enthalten, durch den die Expression der Nukleinsäuresequenzen im erfindungsgemäßen Nukleinsäurekonstrukt in den Organismen vorteilhaft in den Pflanzen zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige vorteilhafte Pflanzenpromotoren sind beispielsweise der *PRP1*-Promotor [Ward et al., *Plant. Mol. Biol.* 22(1993), 361-366], ein durch Benzensulfonamid-induzierbarer (EP 388186), ein durch Tetracyclin-induzierbarer (Gatz et al., (1992) *Plant J.* 2, 397-404), ein durch Salizylsäure induzierbarer Promotor (WO 95/19443), ein durch Abscisinsäure-induzierbarer (EP 335528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer (WO 93/21334) Promotor. Weitere Pflanzenpromotoren sind beispielsweise der Promotor der cytosolischen *FBPase* aus Kartoffel, der *ST-LSI* Promotor aus Kartoffel (Stockhaus et al., *EMBO J.* 8 (1989) 2445-245), der Promotor der Phospho-

ribosylpyrophosphat-Amidotransferase aus *Glycine max* (siehe auch Genbank Accession Nummer U87999) oder ein Nodien-spezifischen Promotor wie in EP 249676 können vorteilhaft verwandt werden.

- 5 In der Expressionskassette (= Genkonstrukt, Nukleinsäurekonstrukt) können wie oben beschrieben noch weitere Gene, die in die Organismen eingebracht werden sollen, enthalten sein. Diese Gene können unter getrennter Regulation oder unter der gleichen Regulationsregion wie die im Verfahren verwendeten Nukleinsäuresequenzen liegen. Bei diesen Genen handelt es sich beispielsweise um Biosynthesegene des Stoff-
- 10 wechselfs wie Gene die an den Stoffwechselwegen der durch die erfindungsgemäßen Nukleinsäuren codierten Proteine beteiligt sind. Es können aber auch Biosynthesegene anderer Stoffwechselwege sein wie der Fettsäure-, der Aminosäure- oder der Vitaminbiosynthese oder Regulationsgene um nur einige zu nennen.
- 15 Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für die erfindungsgemäße Expressionskassette und das erfindungsgemäße Verfahren, wie unten beschrieben, verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.
- 20 Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente (= erfindungsgemäße Nukleinsäuren) miteinander können an die Fragmente Adaptoren oder Linker angesetzt werden.
- 25 Zweckmäßigerweise können die Promotor- und die Terminator-Regionen in Transkriptionsrichtung mit einem Linker oder Polylinker, der eine oder mehrere Restriktionsstellen für die Insertion dieser Sequenz enthält, versehen werden. In der Regel hat der Linker 1 bis 10, meistens 1 bis 8, vorzugsweise 2 bis 6 Restriktionsstellen. Im allge-
- 30 meinen hat der Linker innerhalb der regulatorischen Bereiche eine Größe von weniger als 100 bp, häufig weniger als 60 bp, mindestens jedoch 5 bp. Der Promotor kann sowohl nativ bzw. homolog als auch fremdartig bzw. heterolog zum Wirtsorganismus beispielsweise zur Wirtspflanze sein. Die Expressionskassette beinhaltet in der 5'-3'-Transkriptionsrichtung den Promotor, eine DNA-Sequenz, die für die im erfindungsgemäßen Verfahren verwendeten Proteine codiert und eine Region für die transkriptionale Termination. Verschiedene Terminationsbereiche können vorteilhaft gegeneinander ausgetauscht werden.
- 35
- 40 Ferner können Manipulationen, die passende Restriktionsschnittstellen bereitstellen oder die überflüssige DNA oder Restriktionsschnittstellen entfernen, eingesetzt

werden. Wo Insertionen, Deletionen oder Substitutionen wie z.B. Transitionen und Transversionen in Frage kommen, können in vitro-Mutagenese, -primerrepair-, Restriktion oder Ligation verwendet werden. Bei geeigneten Manipulationen, wie z.B. Restriktion, -chewing-back- oder Auffüllen von Überhängen für -bluntends-, können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden.

Von Bedeutung für eine vorteilhafte hohe Expression kann u.a. das Anhängen des spezifischen ER-Retentionssignals SEKDEL sein (Schouten, A. et al., Plant Mol. Biol. 30 (1996), 781-792), die durchschnittliche Expressionshöhe wird damit verdreifacht bis vervierfacht. Es können auch andere Retentionssignale, die natürlicherweise bei im ER lokalisierten pflanzlichen und tierischen Proteinen vorkommen, für den Aufbau der Kassette eingesetzt werden.

Bevorzugte Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA-Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACH5 entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder entsprechende funktionelle Äquivalente.

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors mit einer geeigneten Nukleinsäuresequenz sowie einem Polyadenylierungssignal nach gängigen Rekombinations- und Klonierungstechniken, wie sie beispielsweise in T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987) beschrieben werden.

Bei der Präparation einer Expressionskassette können verschiedene DNA-Fragmente manipuliert werden, um eine Nukleotid-Sequenz zu erhalten, die zweckmäßigerweise in der korrekten Richtung liest und die mit einem korrekten Leseraster ausgestattet ist. Für die Verbindung der DNA-Fragmente miteinander können an die Fragmente Adaptoren oder Linker angesetzt werden.

Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen beinhalten alle Sequenzmerkmale, die notwendig sind, um eine für den Ort der biologischen Wirkung bzw. Aktivität korrekte Lokalisation zu erreichen. Daher sind keine weiteren Targetingsequenzen per se notwendig. Allerdings kann eine solche Lokalisation wünschenswert und vorteilhaft sein und daher künstlich verändert oder verstärkt

werden, so dass solche Fusionskonstrukte eine bevorzugte vorteilhafte Ausführungsform der Erfindung sind.

Vorteilhaft sind dafür beispielsweise Sequenzen, die ein Targeting in Plastiden gewährleisten. Unter bestimmten Umständen kann auch ein Targeting in andere Kompartimente (referiert: Kermode, Crit. Rev. Plant Sci. 15, 4 (1996), 285-423) z.B. in die Vakuole, in das Mitochondrium, in das Endoplasmatische Retikulum (ER), Peroxisomen, Lipidkörper oder durch ein Fehlen entsprechender operativer Sequenzen ein Verbleib im Kompartiment des Entstehens, dem Zytosol, wünschenswert sein.

Vorteilhafterweise werden die erfindungsgemäßen Nukleinsäuresequenzen zusammen mit mindestens einem Reportergen in eine Expressionskassette kloniert, die in den Organismus über einen Vektor oder direkt in das Genom eingebracht wird. Dieses Reportergen sollte eine leichte Detektierbarkeit über einen Wachstums-, Fluoreszenz-, Chemo-, Biolumineszenz- oder Resistenzassay oder über eine photometrische Messung ermöglichen. Beispielhaft seien als Reportergene Antibiotika- oder Herbizidresistenzgene, Hydrolasegene, Fluoreszenzproteingene, Biolumineszenzgene, Zucker- oder Nukleotidstoffwechselgene oder Biosynthesegene wie das Ura3-Gen, das *llv2*-Gen, das Luciferasegen, das β -Galactosidasegen, das *gfp*-Gen, das 2-Desoxyglucose-6-phosphat-Phosphatasegen, das β -Glucuronidase-Gen, β -Lactamasegen, das Neomycinphosphotransferasegen, das Hygromycinphosphotransferasegen oder das BASTA (= Gluphosinatresistenz)-Gen genannt. Weitere vorteilhafte Antibiotika- oder Herbizidresistenzen sind Resistenzen gegen z.B. Imidazolinon oder Sulfonylharnstoff; die Antibiotikaresistenzen gegen z.B. Bleomycin, Streptomycin, Kanamycin, Tetracyclin, Chloramphenicol, Gentamycin, Geneticin (G418), Spectinomycin oder Blasticidin um nur einige zu nennen. Diese Gene ermöglichen eine leichte Messbarkeit und Quantifizierbarkeit der Transcriptionsaktivität und damit der Expression der Gene. Damit lassen sich Genomstellen identifizieren, die eine unterschiedliche Produktivität zeigen.

Gemäß einer bevorzugten Ausführungsform umfasst eine Expressionskassette stromaufwärts, d.h. am 5'-Ende der codierenden Sequenz, einen Promotor und stromabwärts, d.h. am 3'-Ende, ein Polyadenylierungssignal und gegebenenfalls weitere regulatorische Elemente, welche mit der dazwischenliegenden codierenden Sequenz für die im erfindungsgemäßen Verfahren verwendeten Proteinen operativ verknüpft sind. Unter einer operativen Verknüpfung versteht man die sequenzielle Anordnung von Promotor, codierender Sequenz, Terminator und ggf. weiterer regulatorischer Elemente derart, dass jedes der regulatorischen Elemente seine Funktion bei der Expression der codierenden Sequenz bestimmungsgemäß erfüllen kann. Die zur operativen Verknüpfung bevorzugten Sequenzen sind Targeting-Sequenzen zur

Gewährleistung der subzellulären Lokalisation in Plastiden. Auch Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Mitochondrium, im Endoplasmatischen Retikulum (= ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten sind bei Bedarf einsetzbar sowie Translationsverstärker wie die 5'-Führungssequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711).

Eine Expressionskassette kann beispielsweise einen konstitutiven Promotor beispielsweise den 35S-, 34S- oder eine Ubiquitin-Promotor, das zu exprimierende Gen und das ER-Retentionssignal enthalten. Als ER-Retentionssignal wird bevorzugt die Aminosäuresequenz KDEL (Lysin, Asparaginsäure, Glutaminsäure, Leucin) verwendet.

Die Expressionskassette wird zur Expression in einem prokaryontischen oder eukaryontischen Wirtsorganismus beispielsweise einem Mikroorganismus wie einem Pilz oder einer Pflanze vorteilhafterweise in einen Vektor wie beispielsweise einem Plasmid, einem Phagen oder sonstiger DNA inseriert, der eine optimale Expression der Gene im Wirtsorganismus ermöglicht. Geeignete Plasmide sind beispielsweise in E. coli pLG338, pACYC184, pBR-Serie wie z.B. pBR322, pUC-Serie wie pUC18 oder pUC19, M113mp-Serie, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III¹¹³-B1, λ gt11 oder pBdCl, in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus pUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667, in Pilzen pALS1, pIL2 oder pBB116, weitere vorteilhafte Pilzvektoren werden von Romanos, M.A. et al., [(1992) "Foreign gene expression in yeast: a review", Yeast 8: 423-488] und von van den Hondel, C.A.M.J.J. et al. [(1991) "Heterologous gene expression in filamentous fungi"] sowie in More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego] und in "Gene transfer systems and vector development for filamentous fungi" [van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) in: Applied Molecular Genetics of Fungi, Peberdy, J.F. et al., eds., p. 1-28, Cambridge University Press: Cambridge] beschrieben. Vorteilhafte Hefepromotoren sind beispielsweise 2 μ M, pAG-1, YEp6, YEp13 oder pEMBLye23. Beispiele für Algen- oder Pflanzenpromotoren sind pLGV23, pGHlac⁺, pBIN19, pAK2004, pVKH oder pDH51 (siehe Schmidt, R. and Willmitzer, L., 1988). Die oben genannten Vektoren oder Derivate der vorstehend genannten Vektoren stellen eine kleine Auswahl der möglichen Plasmide dar. Weitere Plasmide sind dem Fachmann wohl bekannt und können beispielsweise aus dem Buch Cloning Vectors (Eds. Pouwels P. H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018) entnommen werden. Geeignete pflanzliche Vektoren werden unter anderem in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Kap. 6/7, S.71-119 beschrieben. Vorteilhafte Vektoren sind sog. shuttle-Vektoren oder binäre Vektoren, die in E. coli und Agrobacterium replizieren.

Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren wie beispielsweise Phagen, Viren wie SV40, CMV, Baculovirus, Adenovirus, Transposons, IS-Elemente, Phasmide, Phagemide, Cosmide, lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder
5 chromosomal repliziert werden bevorzugt ist eine chromosomale Replikation. Umfasst sind funktionelle und nicht-funktionelle Vektoren.

In einer weiteren Ausgestaltungsform des Vektors kann die erfindungsgemäße
10 Nukleinsäurekonstrukt auch vorteilhafterweise in Form einer linearen DNA in die Organismen eingeführt werden und über heterologe oder homologe Rekombination in das Genom des Wirtsorganismus integriert werden. Diese lineare DNA kann aus einem linearisierten Plasmid oder nur aus dem Nukleinsäurekonstrukt als Vektor oder den verwendeten Nukleinsäuresequenzen bestehen.

15 In einer weiteren vorteilhaften Ausführungsform können die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen auch alleine in einen Organismus eingebracht werden.

20 Sollen neben den Nukleinsäuresequenzen weitere Gene in den Organismus eingeführt werden, so können alle zusammen mit einem Reportergen in einem einzigen Vektor oder jedes einzelne Gen mit oder ohne einem Reportergen in je einem Vektor in den Organismus eingebracht werden, wobei die verschiedenen Vektoren gleichzeitig oder sukzessive eingebracht werden können.

25 Der Vektor enthält vorteilhaft mindestens eine Kopie der verwendeten Nukleinsäuresequenzen und/oder des erfindungsgemäßen Nukleinsäurekonstrukts.

30 Beispielhaft kann das Nukleinsäurekonstrukt in den Tabak-Transformationsvektor pBinAR eingebaut werden und unter der Kontrolle des 35S-, 34S- oder Ubiquitin-Promotor oder des USP-Promotor stehen.

Alternativ kann ein rekombinanter Vektor (= Expressionsvektor) auch in-vitro transkribiert und translatiert werden, z.B. durch Nutzung des T7 Promotors und der T7 RNA
35 Polymerase.

Weitere vorteilhafte Vektoren enthalten in Pflanzen oder Pflanzenkulturen nutzbare Resistenzen wie die Phosphinotricin- (= bar-Resistenz), die Methioninsulfoximin-, die Sulfonylharnstoff- (= ilv-Resistenz, ind *S. cerevisiae* ilv2), die Phenoxyphenoxyherbizid-
40 (= ACCase-Resistenz), die Glyphosate- oder Clearfieldresistenz (AHAS-Resistenz)

bzw. die Gene, die für diese Resistenzen codieren. Diese Resistenzen sind in ganzen Pflanzen zur Selektion von transgenen Pflanzen nutzbar. Nur Pflanzen, die über eine Transformation, diese Resistenzen erhalten haben, können in Gegenwart der selektierenden Substanz wachsen. In Zellkulturen auf Agarplatten können nach Transformation in planta - beispielsweise Infiltration der Samenvorläuferzellen - auch beispielsweise Kanamycin oder Hygromycin als selektierendes Agens verwendet werden. Darüber hinaus können vorteilhafte Vektoren Sequenzen für die Integration in das Genom der Organismen bevorzugt der Pflanzen enthalten. Beispiele für derartige Sequenzen sind die sog. T-DNA-Borders. Außerdem können vorteilhafte Vektoren noch Promotoren und Terminatoren, wie beispielsweise die oben beschriebenen, enthalten. Auch sogenannte PolyA-Sequenzen können im Vektor enthalten sein. Vorteileilhafte Vektoren sind beispielsweise Figur 1, 2 und 3 zu entnehmen. SEQ ID NO: 25 gibt die vorteilhafte Sequenz des Vektors pMTX 1a300 wieder. Dieser enthält eine Kanamycin-Resistenz (Nukleotid 4922-5713), eine Phosphinotricin-Resistenz (Nukleotid 6722-7288), das LacZalpha-Fragment (Nukleotid 7630-7864), einen Teil von pVS1sta (Nukleotid 945-1945), einen Teil von pBR322bom (Nukleotid 3948-4208), eine T-Border-Sequenz (links, Nukleotid 6138-6163), eine T-Border-Sequenz (rechts, Nukleotid 7924-7949), ein PolyA-Teil (Nukleotid 7292 - 7503), den mas2'1'-Promotor (Nukleotid 6241-6718) und zwei Replikations-Origins pVS1rep (Nukleotid 6241-6718) sowie pBR322ori (Nukleotid 43-4628).

In Prokaryoten verwendete Expressionsvektoren nutzen häufig induzierbare Systeme mit und ohne Fusionsproteinen bzw Fusionsoligopeptiden, wobei diese Fusionen sowohl N-terminal als auch C-terminal oder anderen nutzbaren Domänen eines Proteins erfolgen können. Solche Fusionsvektoren dienen in der Regel dazu: i.) die Expressionsrate der RNA zu erhöhen ii.) die erzielbare Proteinsyntheserate zu erhöhen, iii.) die Löslichkeit des Proteins zu erhöhen, iv.) oder die Reinigung durch einen für die Affinitätschromatographie nutzbare Bindesequenz zu vereinfachen. Häufig werden auch proteolytische Spaltstellen über Fusionsproteine eingeführt, was die Abspaltung eines Teils des Fusionsproteins auch der Reinigung ermöglicht. Solche Erkennungssequenzen für Proteasen erkennen sind z.B. Faktor Xa, Thrombin und Enterokinase.

Typische vorteilhafte Fusions- und Expressionsvektoren sind pGEX [Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) Gene 67:31-40], pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) welches Glutathion S-transferase beinhaltet (GST), Maltose Bindeprotein, oder Protein A.

Weitere Beispiele für E. coli Expressionsvektoren sind pTrc [Amann et al., (1988) Gene 69:301-315] und pET Vektoren [Studier et al., Gene Expression Technology: Methods

in Enzymology Academic Press, San Diego, California () 60-89; Stratagene, Amsterdam, Niederlande].

Weitere vorteilhafte Vektoren zur Verwendung in Hefe sind pYepSec1 (Baldari, et al.,
5 (1987) Embo J. 6:229-234), pMFa (Kurjan and Herskowitz, (1982) Cell 30:933-943),
pJRY88 (Schultz et al., (1987) Gene 54:113-123), and pYES-Derivate (Invitrogen
Corporation, San Diego, CA). Vektoren für die Nutzung in filamentösen Pilzen sind
beschrieben in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer
10 systems and vector development for filamentous fungi, in: Applied Molecular Genetics
of Fungi, J.F. Peberdy, et al., eds., p. 1-28, Cambridge University Press: Cambridge.

Alternativ können auch vorteilhaft Insektenzelleexpressionsvektoren genutzt werden z.B.
für die Expression in Sf 9 Zellen. Dies sind z.B. die Vektoren der pAc Serie (Smith et al.
(1983) Mol. Cell Biol. 3:2156-2165) und der pVL series (Lucklow and Summers (1989)
15 Virology 170:31-39).

Des weiteren können zur Genexpression vorteilhaft Pflanzenzellen oder Algenzellen
genutzt werden. Beispiele für Pflanzenexpressionsvektoren finden sich in Becker, D.,
et al. (1992) "New plant binary vectors with selectable markers located proximal to
20 the left border", Plant Mol. Biol. 20: 1195-1197 oder in Bevan, M.W. (1984) "Binary
Agrobacterium vectors for plant transformation", Nucl. Acid. Res. 12: 8711-8721.

Weiterhin können die erfindungsgemäßen Nukleinsäuresequenzen in Säugerzellen
exprimiert werden. Beispiel für entsprechende Expressionsvektoren sind pCDM8 und
25 pMT2PC genannt in: Seed, B. (1987) Nature 329:840 oder Kaufman et al. (1987)
EMBO J. 6:187-195). Dabei sind vorzugsweise zu nutzende Promotoren viralen
Ursprungs wie z.B. Promotoren des Polyoma, Adenovirus 2, Cytomegalovirus oder
Simian Virus 40. Weitere prokaryotische und eukaryotische Expressionssysteme sind
genannt in Kapitel 16 und 17 in Sambrook et al., Molecular Cloning: A Laboratory
30 Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory
Press, Cold Spring Harbor, NY, 1989. Weitere vorteilhafte Vektoren werden in Hellens
et al. (Trends in plant science, 5, 2000) beschrieben.

Das Einbringen der erfindungsgemäßen Nukleinsäuren, der Expressionskassette oder
35 des Vektors in Organismen beispielsweise in Pflanzen kann prinzipiell nach allen dem
Fachmann bekannten Methoden erfolgen.

Für Mikroorganismen kann der Fachmann entsprechende Methoden den Lehrbüchern
von Sambrook, J. et al. (1989) Molecular cloning: A laboratory manual, Cold Spring
40 Harbor Laboratory Press, von F.M. Ausubel et al. (1994) Current protocols in molecular

biology, John Wiley and Sons, von D.M. Glover et al., DNA Cloning Vol.1, (1995), IRL Press (ISBN 019-503476-9), von Kaiser et al. (1994) Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press oder Guthrie et al. Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, 1994, Academic Press entnehmen.

5

Die Übertragung von Fremdgenen in das Genom einer Pflanze wird als Transformation bezeichnet. Es werden dabei die beschriebenen Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengewebe oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt. Geeignete Methoden sind die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone - die sogenannte particle bombardment Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Mikroinjektion und der durch Agrobacterium vermittelte Gentransfer vorteilhaft erfolgt der Gentransfer in der vorliegenden Erfindung mit beispielsweise dem Agrobacterium tumefaciens Stamm GV 3101 pMP90. Die genannten Verfahren sind beispielsweise in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press (1993) 128-143 sowie in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225) beschrieben. Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, Agrobacterium tumefaciens zu transformieren, beispielsweise pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Mit einem solchen Vektor transformierte Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen, wie z.B. von Tabakpflanzen, verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden. Die Transformation von Pflanzen mit Agrobacterium tumefaciens wird beispielsweise von Höfgen und Willmitzer in Nucl. Acid Res. (1988) 16, 9877 beschrieben oder ist unter anderem bekannt aus F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, 1993, S. 15-38.

Im folgenden ist eine vorteilhafte Ausführungsform wiedergegeben. Werden für die Transformation Agrobakterien verwendet, wird die einzuführende Nukleinsäure bzw. DNA in spezielle Plasmide cloniert, und zwar entweder in einen intermediären Vektor oder in einen binären Vektor. Die intermediären Vektoren können aufgrund von Sequenzen, die homolog zu Sequenzen in der T-DNA sind, durch homologe Rekombination in das Ti- oder Ri-Plasmid der Agrobakterien integriert werden. Dieses enthält außerdem die für den Transfer der T-DNA notwendige vir-Region. Intermediäre Vektoren können nicht in Agrobakterien replizieren. Mittels eines Helferplasmids kann der intermediäre Vektor auf Agrobacterium tumefaciens übertragen werden (Konju-

40

- gation). Binäre Vektoren können sowohl in *E. coli* als auch in *Agrobacterium* replizieren. Sie enthalten ein Selektionsmarker-Gen und einen Linker oder Polylinker, welche von der rechten und linken T-DNA Grenzregion eingerahmt werden. Sie können direkt in die *Agrobacterium* transformiert werden (Holsters et al. Mol. Gen. Genet. 163 (1978), 181-187). Das als Wirtszelle dienende *Agrobacterium* soll ein Plasmid, das eine vir-Region trägt, enthalten. Die vir-Region ist für den Transfer der T-DNA in die Pflanzenzelle notwendig. Zusätzliche T-DNA kann vorhanden sein. Das derartig transformierte *Agrobacterium* wird zur Transformation von Pflanzenzellen verwendet.
- 10 Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in EPA-0 120 516; Hoekema, In: The Binary Plant Vector System Offsetdrukkerij Kanters B.V., Alblasterdam (1985), Chapter V; Fraley et al., Crit. Rev. Plant. Sci., 4: 1-46 und An et al. EMBO J. 4 (1985), 277-287 beschrieben worden.
- 15 Für den Transfer der DNA in die Pflanzenzelle können Pflanzen-Explantate zweckmäßigerweise mit *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* kokultiviert werden. Aus dem infizierten Pflanzenmaterial (z.B. Blattstücke, Stengel-segmente, Wurzeln, aber auch Protoplasten oder Suspensions-kultivierte Pflanzen-
- 20 zellen) können dann in einem geeigneten Medium, welches Antibiotika oder Biozide zur Selektion transformierter Zellen enthalten kann, wieder ganze Pflanzen regeneriert werden. Die so erhaltenen Pflanzen können dann auf Anwesenheit der eingeführten DNA untersucht werden. Andere Möglichkeiten der Einführung fremder DNA unter Verwendung des biolistischen Verfahrens oder durch Protoplastentransformation sind
- 25 bekannt (vergl. z.B. Willmitzer, L., 1993 Transgenic plants. In: Biotechnology, A Multi-Volume Comprehensive Treatise (H.J. Rehm, G. Reed, A. Pühler, P. Stadler, eds.), Vol. 2, 627-659, VCH Weinheim-New York-Basel-Cambridge.).
- 30 Auch die Transformation monokotyler Pflanzen mittels *Agrobacterium* basierender Vektoren wurde beschrieben (Chan et al, Plant Mol. Biol. 22(1993), 491-506; Hiei et al, Plant J. 6 (1994) 271-282; Deng et al; Science in China 33 (1990), 28-34; Wilmink et al, Plant Cell Reports 11,(1992) 76-80; May et al; Biotechnology 13 (1995) 486-492; Conner und Domisse; Int. J. Plant Sci. 153 (1992) 550-555; Ritchie et al; Transgenic Res. (1993) 252-265). Alternative Systeme zur Transformation von monokotylen
- 35 Pflanzen sind die Transformation mittels des biolistischen Ansatzes (Wan und Lemaux; Plant Physiol. 104 (1994), 37-48; Vasil et al; Biotechnology 11 (1992), 667-674; Ritala et al, Plant Mol. Biol 24, (1994) 317-325; Spencer et al, Theor. Appl. Genet. 79 (1990), 625-631) die Protoplastentransformation, die Elektroporation von partiell permeabilisierten Zellen; die Einbringung von DNA mittels Glasfasern. Insbesondere die
- 40 Transformation von Mais wurde in der Literatur mehrfach beschrieben (vgl. z.B.

WO 95/06128; EP 013849 A1; EP 0465875 A1; EP 0292435 A1; Fromm et al, Biotechnology 8 (1990), 833-844; Gordon-Kamm et al, Plant Cell 2 (1990), 603-618; Koziel et al, Biotechnology 11(1993) 194-200; Moroc et al, Theor Applied Genetics 80 (1990) 721-726).

5

Auch die erfolgreiche Transformation anderer Getreidearten wurde bereits beschrieben z.B. für Gerste (Wan und Lemaux, s.o.; Ritala et al, s.o.; Weizen (Nehra et al, Plant J. 5(1994) 285-297).

- 10 Mit einem erfindungsgemäßen Vektor transformierte Agrobakterien können ebenfalls in bekannter Weise zur Transformation von Pflanzen wie Testpflanzen wie Arabidopsis oder Kulturpflanzen wie Getreide, Mais, Hafer, Roggen, Gerste, Weizen, Soja, Reis, Baumwolle, Zuckerrübe, Canola, Sonnenblume, Flachs, Hanf, Kartoffel, Tabak, Tomate, Karotte, Paprika, Raps, Tapioka, Maniok, Pfeilwurz, Tagetes, Alfalfa, Salat
- 15 und den verschiedenen Baum-, Nuss- und Weinspezies verwendet werden, z.B. indem verwundete Blätter oder Blattstücke in einer Agrobakterienlösung gebadet und anschließend in geeigneten Medien kultiviert werden.

- Die genetisch veränderten Pflanzenzellen können über alle dem Fachmann bekannten
- 20 Methoden regeneriert werden. Entsprechende Methoden können den oben genannten Schriften von S.D. Kung und R. Wu, Potrykus oder Höfgen und Willmitzer entnommen werden.

- Unter Pflanzen im Sinne der Erfindung sind Pflanzenzellen, -gewebe, -organe oder
- 25 ganzen Pflanzen wie Samen, Knollen, Blüten, Pollen, Früchte, Sämlinge, Wurzeln, Blätter, Stengel oder sonstige Pflanzenteile zu verstehen. Außerdem ist unter Pflanzen Vermehrungsmaterial wie Samen, Früchte, Sämlinge, Stecklinge, Knollen, Schnitte oder Wurzelstöcke zu verstehen.

- 30 Als Organismen bzw. Wirtsorganismen für die erfindungsgemäße Nukleinsäure, die Expressionskassette oder den Vektor eignen sich prinzipiell vorteilhaft alle Organismen, die in der Lage sind die erfindungsgemäß verwendeten Nukleinsäuren zu exprimieren bzw. für die Expression rekombinanter Gene geeignet sind. Beispielhaft seien Pflanzen wie Arabidopsis, Asteraceae wie Calendula oder Kulturpflanzen wie
- 35 Soja, Erdnuss, Rizinus, Sonnenblume, Mais, Baumwolle, Flachs, Raps, Kokosnuss, Ölpalme, Färbersaflor (*Carthamus tinctorius*) oder Kakaobohne, Mikroorganismen wie Pilze beispielsweise die Gattung *Mortierella*, *Saprolegnia* oder *Pythium*, Bakterien wie die Gattung *Escherichia*, Hefen wie die Gattung *Saccharomyces*, Cyanobakterien, Ciliaten, Algen oder Protozoen wie Dinoflagellaten wie *Cryptocodinium* genannt.
- 40 Beispielsweise seien Organismen, die natürlicherweise Öle in größeren Mengen

synthetisieren wie Soja, Raps, Kokosnuss, Ölpalme, Färbers Rizinus, Calendula, Erdnuss, Kakaobohne oder Sonnenblume genannt. Prinzipiell sind als Wirtsorganismen auch nicht-humane transgene Tiere geeignet beispielsweise *C. elegans*.

5 Bevorzugt sind transgene Pflanzen, die ein erfindungsgemäßes funktionelles oder nicht funktionelles Nukleinsäurekonstrukt oder einen erfindungsgemäßen funktionellen oder nicht funktionellen Vektor enthalten. Unter funktionell ist im Sinne der Erfindung zu verstehen, dass die im Verfahren verwendeten Nukleinsäuren allein oder im Nukleinsäurekonstrukt oder im Vektor exprimiert werden und ein biologisch aktives Gen-

10 produkt hergestellt wird. Unter nicht funktionell ist im Sinne der Erfindung zu verstehen, dass die im Verfahren verwendeten Nukleinsäuren allein oder im Nukleinsäurekonstrukt oder im Vektor nicht transkribiert, nicht exprimiert werden und/oder ein biologisch inaktives Genprodukt hergestellt wird. In diesem Sinne handelt es sich bei den sogenannten Antisense-RNAs auch um nicht funktionelle Nukleinsäuren bzw. bei Insertion

15 in das Nukleinsäurekonstrukt oder den Vektor um ein nicht funktionelles Nukleinsäurekonstrukt oder nicht funktionellen Vektor. Sowohl das erfindungsgemäße Nukleinsäurekonstrukt als auch der erfindungsgemäße Vektor kann zur Herstellung von transgenen Organismen bevorzugt Pflanzen vorteilhaft verwendet werden.

20 Unter transgen im Sinne der Erfindung ist zu verstehen, dass die im Verfahren verwendeten Nukleinsäuren nicht an ihrer natürlichen Stelle im Genom eines Organismus sind, dabei können die Nukleinsäuren homolog oder heterolog exprimiert werden. Transgen bedeutet aber auch, dass die erfindungsgemäßen Nukleinsäuren an ihrem natürlichen Platz im Genom eines Organismus sind, dass jedoch die Sequenz gegenüber der natürlichen Sequenz verändert wurde und/oder dass die Regulationssequenzen, der natürlichen Sequenzen verändert wurden. Bevorzugt ist unter transgen die Expression der Nukleinsäuren an nicht natürlicher Stelle im Genom zu verstehen, das heißt eine homologe oder bevorzugt heterologe Expression der Nukleinsäuren liegt vor. Gleiches gilt für das erfindungsgemäße Nukleinsäurekonstrukt oder den Vektor.

30 Nutzbare Wirtszellen sind weiterhin genannt in: Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

Verwendbare Expressionsstämme z.B. solche, die eine geringere Proteaseaktivität aufweisen sind beschrieben in: Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128.

35

Weiterhin umfasst die Erfindung auch die Verwendung der erfindungsgemäßen Nukleinsäuren, z.B. der unter den SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,

40

SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargelegten Nukleotidsequenzen zur Erstellung von genetisch veränderten Pflanzen, die dadurch gekennzeichnet sind, dass sie veränderte Proteine, der von den SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 codierten Proteine enthalten, die dadurch gekennzeichnet sind dass sie eine sehr viele geringere Interaktion mit dem Herbizid aufweisen bzw. in ihrer Aktivität nicht durch das Herbizid beeinträchtigt werden.

Die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, deren aufgrund des degenerierten genetischen Codes abgeleiteten Sequenzen sowie deren Derivate wurden aus einer Populationen transgener Pflanzen identifiziert, die zum einen dadurch gekennzeichnet war, dass sie mittels des Agrobacterium transformiert worden sind und im Rahmen dieses Prozesses neue DNA im Chromosomen an zufälliger Stelle integriert worden war. Durch Rückkreuzungen wurden schließlich Pflanzen isoliert, die die identifizierten Nukleinsäuren auf beiden homologen Chromosomen enthalten. Diese Pflanzen sind lethal, daher sterben sie entweder schon als Embryo oder aber als Keimling. Es konnten keine homozygoten Linien gewonnen werden. Darüber hinaus sind diese Pflanzen dadurch gekennzeichnet, dass sie im Verlauf des Screenings als für lethale Mutationen segregierende Linien identifiziert worden sind. Diese Pflanzen weisen als Ergebnis der homozygoten Situation der Integration der neuen DNA starke Behinderungen im Wachstum und oder Entwicklung auf. Es ist davon auszugehen, dass diese Wachstums- und Entwicklungsbehinderungen darauf zurückzuführen sind, dass die neu inserierte DNA in für Wachstum und Entwicklung wichtige Gene integriert hat, wodurch in der homozygoten Situation deren biologische Funktion eingeschränkt oder blockiert wird. Dieses bedeutet, dass diese Gene sowie deren aufgrund des degenerierten genetischen Codes abgeleiteten Sequenzen sowie deren Derivate für Proteine codieren, welche in analoger Weise wie für die in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5,

SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,
SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25,
SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35,
SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45,
5 SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 beschrieben, geeignete Ziel-
proteine für neu zu entwickelnde Herbizide darstellen.

Zur Erzeugung veränderter Proteine werden in einer vorteilhaften Ausführungsform
die genannten Nukleinsäuren überexprimiert und folgende Prozessschritte werden
10 vorteilhaft durchlaufen:

a) Expression der von den SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5,
SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,
SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23,
15 SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31,
SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,
SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
SEQ ID NO: 49 oder SEQ ID NO: 51 oder von einer Nukleinsäuresequenz, die
sich aufgrund des degenerierten genetischen Codes aus den durch Rücküber-
20 setzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8,
SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16,
SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24,
SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32,
SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40,
25 SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48,
SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen
ableiten lässt, oder von Derivaten oder Fragmenten der in SEQ ID NO: 1,
SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11,
SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19,
30 SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27,
SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35,
SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43,
SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51
dargestellten Nukleinsäuresequenzen, die für Polypeptide mit den in
35 SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10,
SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18,
SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26,
SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34,
SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42,
40 SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder

SEQ ID NO: 3 dargestellten Aminosäuresequenzen codieren und mindestens 50 %, 60 %, vorzugsweise 70 %, 80 %, 90 % oder mehr Homologie auf Aminosäureebene aufweisen, codierten Proteine in einem heterologen System beispielsweise einem Mikroorganismus wie einem Bakterium der Gattung Escherichia wie E. coli XL1-Red oder in einem zellfreien System.

5

b) Randomisierte oder gerichtete Mutagenese des Proteins durch Modifikation der Nukleinsäure,

10

c) Messung der Interaktion oder der biologischen Aktivität des veränderten Proteins mit dem Herbizid bzw. in Gegenwart des Herbizids,

d) Identifizierung von Derivaten des Proteins die eine geringere Interaktion oder eine wenig beeinflusste biologische Aktivität aufweisen,

15

e) Testung der biologischen Aktivität des Proteins nach Applikation des Herbizides.

Vorteilhaft erfolgt die Überführung des so erhaltenen veränderten Proteins bzw. der veränderten Nukleinsäure, z.B. der unter SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 aufgeführten Sequenzen sowie der weiteren oben beschriebenen erfindungsgemäßen Sequenzen, z.B. Derivate und Fragmente, z.B. von anderen Pflanzen, in einen Organismus vorteilhaft in eine Pflanze, bevorzugt pflanzliche Zellen.

20

25

Eine weitere Ausführungsform der Erfindung ist ein Verfahren zur Erzeugung veränderter Genprodukte, die von den hierin beschriebenen erfindungsgemäßen Nukleinsäuresequenzen, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 codiert werden, dadurch gekennzeichnet, dass es folgende Prozessschritte umfasst:

30

35

a) Expression der von den SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15,

40

- 5 SEQ ID NO: 7, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23,
SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31,
SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39,
SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47,
SEQ ID NO: 49 oder SEQ ID NO: 51 codierten Proteine oder Derivate oder
Fragmente, z.B. von anderen Pflanzen, davon in einem heterologen System o-
der in einem Zellfreisystem
- 10 b) Randomisierte oder gerichtete Mutagenese des Proteins durch Modifikation der
Nukleinsäure,
- 15 c) Messung der Interaktion des veränderten Genprodukts mit dem Herbizid oder
der biologischen Aktivität des veränderten Genprodukts in Gegenwart des
Herbizids,
- d) Identifizierung von Derivaten des Proteins die eine geringere Interaktion auf-
weisen oder in ihrer Aktivität weniger beeinflusst sind,
- 20 e) Testung der biologischen Aktivität des Proteins nach Applikation des Herbizides,
- f) Auswahl der Nukleinsäuresequenzen, die oder deren Genprodukte eine ver-
änderte biologische Aktivität gegenüber dem Herbizid, vorzugsweise eine verrin-
gerte Hemmung durch das Herbizid oder geringere Interaktion mit dem Herbizid,
aufweisen.
- 25 Die nach dem oben beschriebenen Verfahren nach ausgewählten Sequenzen
können vorteilhaft in einen Organismus eingebracht werden. Deshalb ist ein weiterer
Erfindungsgegenstand ein nach diesem Verfahren hergestellter Organismus, bevor-
zugt ist der Organismus eine Pflanze. Das Verfahren eignet sich auch für die Gen-
30 expression der oben genannten biologisch aktiven Derivate und Fragmente.
- Anschließend erfolgt die Regeneration ganzer Pflanzen und Überprüfung der
Resistenz gegenüber dem Herbizid in intakten Pflanzen.
- 35 Veränderte Proteine und/oder Nukleinsäuren, die in Pflanzen Resistenz gegen
Herbizide vermitteln können, können aus den hierin beschriebenen, erfindungs-
gemäßen Sequenzen, insbesondere aus den Sequenzen SEQ ID NO: 1, SEQ ID
NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13,
SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23,
40 SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33,

SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 oder deren Derivaten aus anderen Pflanzen auch über die sogenannte "site directed mutagenesis" hergestellt werden. Durch diese Mutagenese kann beispielsweise die Stabilität und/oder enzymatische Aktivität von Enzymen oder die Eigenschaften wie Bindung von niedermolekularen Verbindungen mit kleiner 1000 Molekulargewicht gezielt verändert und vorteilhaft verringert werden. Vorteilhaft sollte das Molekulargewicht der Verbindungen kleiner 900 Dalton, bevorzugt kleiner 800, besonders bevorzugt kleiner 700, ganz besonders bevorzugt kleiner 600 Dalton sein, vorteilhaft mit einem Ki-Wert von kleiner 10^{-7} , vorteilhaft kleiner 10^{-8} , bevorzugt kleiner 10^{-9} M, vorteilhaft sollte dabei diese Hemmwirkung auf eine spezifische Hemmung der biologischen Aktivität der erfindungsgemäßen Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine zurückzuführen sein, das heißt es sollte keine Hemmung durch diese niedermolekularen Substanzen weiterer nahe verwandter Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine erfolgen. Weiterhin sollten die niedermolekulare Substanzen vorteilhaft ein Molekulargewicht von größer 50 Dalton, bevorzugt größer 100 Dalton, besonders bevorzugt größer 150 Dalton, ganz besonders bevorzugt größer 200 Dalton haben. Vorteilhaft sollten die niedermolekularen Substanzen weniger als drei Hydroxylgruppen an einem Kohlenstoffatom-enthaltenden Ring haben. Weiterhin sollten auch keine freie(n) Säure- oder Lacton-Gruppe(n) sowie keine Phosphatgruppe und nicht mehr als eine Aminogruppe im Molekül enthalten sein. Auch Basen wie Adenosin im Molekül sind weniger bevorzugt. Auch die Stabilität und/oder enzymatische Aktivität von Enzymen oder die Eigenschaften wie Bindung von Proteinen oder Antisense-RNA kann so sehr gezielt verbessern bzw. verändert werden.

Weiterhin können Veränderungen über die von Spee et al. (Nucleic Acids Research, Vol. 21, No. 3, 1993: 777-78) beschriebenen PCR-Methode unter Verwendung von dITP zur zufälligen Mutagenese erzielt werden oder durch die von Rellos et al. (Protein Expr. Purif., 5, 1994: 270-277) weiter verbesserte Methode.

Eine weitere Möglichkeit zur Herstellung dieser veränderten Proteine und/oder von Nukleinsäuren ist eine von Stemmer et al. (Proc. Natl. Acad. Sci. USA, Vol. 91, 1994: 10747-10751) beschriebene "in vitro" Rekombinationstechnik für die molekulare Evolution oder die von Moore et al. (Nature Biotechnology Vol. 14, 1996: 458-467) beschriebene Kombination der PCR- und Rekombinationsmethode.

Ein weiterer Weg zur Mutagenese von Nukleinsäuren bzw. Proteinen wird von Greener et al. in Methods in Molecular Biology (Vol. 57, 1996: 375-385) beschrieben. In EP-A-0 909 821 wird eine Methode zur Veränderung von Proteinen unter Ver-

wendung des Mikroorganismus *E. coli* XL-1 Red beschrieben. Dieser Mikroorganismus erzeugt bei der Replikation Mutationen in den eingeführten Nukleinsäuren und führt so zu einer Veränderung der genetischen Information. Über Isolierung der veränderten Nukleinsäuren bzw. der veränderten Proteine und Testung auf Resistenz lassen sich leicht vorteilhafte Nukleinsäuren und die durch sie codierten Proteine und umgekehrt identifizieren. Diese können dann nach Einbringen in Pflanzen dort die Resistenz ausprägen und so zur Resistenz gegen die Herbizide führen.

Weitere Methoden der Mutagenese und Selektion sind beispielsweise Methoden wie die in vivo Mutagenese von Samen oder Pollen und Selektion resistenter Allele in Anwesenheit der erfindungsgemäßen Inhibitoren, gefolgt von genetischer und molekularer Identifizierung des veränderten, resistenten Allels. Weiterhin die Mutagenese und Selektion von Resistenzen in der Zellkultur durch Vermehrung der Kultur in Anwesenheit von sukzessiv steigenden Konzentrationen der erfindungsgemäßen Inhibitoren. Dabei kann die Erhöhung der spontanen Mutationsrate durch chemische/physikalische mutagene Behandlung ausgenutzt werden. Wie vorgehend beschrieben lassen sich auch mit Mikroorganismen, die eine endogene oder rekombinante Aktivität der durch die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren codierten Proteine haben, und die gegenüber den erfindungsgemäß identifizierten Inhibitoren sensitiv sind, veränderte Gene isolieren. Die Anzucht der Mikroorganismen auf Medien mit steigenden Konzentrationen von erfindungsgemäßen Inhibitoren erlaubt die Selektion und Evolution von resistenten Varianten der erfindungsgemäßen Targets. Die Frequenz der Mutationen kann wiederum durch mutagenen Behandlungen erhöht werden.

Daneben stehen Verfahren zur gezielten Veränderungen von Nukleinsäuren zur Verfügung (Zhu et al. Proc. Natl. Acad. Sci. USA, Vol. 96, 8768-8773 und Beethem et al., Proc. Natl. Acad. Sci. USA, Vol 96, 8774-8778). Diese Methoden ermöglichen es, in den Proteinen solche Aminosäuren, die für die Bindung von Inhibitoren von Bedeutung sind, durch funktionell äquivalente Aminosäuren zu ersetzen, die jedoch die Bindung des Inhibitors verhindern.

Ein weiterer Erfindungsgegenstand ist weiterhin ein Verfahren zur Erstellung von Nukleotidsequenzen welche für Genprodukte codieren, die eine veränderte biologische Aktivität aufweisen, wobei die biologische Aktivität dahingegen verändert wurde, dass eine erhöhte Aktivität vorliegt. Unter erhöhter Aktivität ist eine gegenüber dem Ausgangsorganismus bzw. gegenüber dem Ausgangsgenprodukt um mindestens 10 %, bevorzugt um mindestens 30 %, besonders bevorzugt um mindestens 50 % oder 70 %, ganze besonders bevorzugt um mindestens 100 % höhere Aktivität zu verstehen. Weiterhin kann die biologische Aktivität dahingegen verändert worden sein, dass die erfindungsgemäßen Substanzen und/oder Mittel nicht mehr oder nicht mehr richtig

- an die Nukleinsäuresequenzen und/oder die durch sie codierten Genprodukte binden. Unter nicht mehr oder nicht mehr richtig ist im Sinne der Erfindung zu verstehen, dass die Substanzen um mindestens 30 %, bevorzugt mindestens 50 %, besonders bevorzugt um mindestens 70 %, ganz besonders bevorzugt um mindestens 80 % oder gar nicht mehr an die veränderten Nukleinsäuren und/oder Genprodukte im Vergleich zum Ausgangsgenprodukt oder den Ausgangsnukleinsäuren binden.

Noch ein weiterer Aspekt der Erfindung betrifft deshalb eine nach dem oben beschriebenen erfindungsgemäßen Verfahren genetisch veränderte transgene Pflanze.

- Genetisch veränderten transgene Pflanzen, die gegen die nach den erfindungsgemäßen Verfahren gefundenen Substanzen und/oder Mittel enthaltend diese Substanzen resistent sind, können auch durch Überexpression der in den erfindungsgemäßen Verfahren verwendeten Nukleinsäuren, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 erzeugt werden.
- Deshalb ist ein weiterer Erfindungsgegenstand ein Verfahren zur Erzeugung transgener Pflanzen, die gegen Substanzen, die nach einem erfindungsgemäßen Verfahren gefunden wurden, resistent sind, dadurch gekennzeichnet, dass in diesen Pflanzen erfindungsgemäße Nukleinsäuren mit einer der beschriebenen biologischen Aktivität, insbesondere mit den Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 überexprimiert werden. Ein ähnliches Verfahren wird beispielhaft in Lemantova et al., Plant Physiol., 122, 2000: 75-83 beschrieben. Natürlich können auch die hierin genannten Derivate und Fragmente, z.B. aus anderen Pflanzen, verwendet werden, die die gewünschte Aktivität aufweisen.
- Die oben beschriebenen erfindungsgemäßen Verfahren zur Erzeugung resistenter Pflanzen ermöglichen die Entwicklung neuer Herbizide, die eine möglichst umfassende Pflanzenspezies unabhängige Wirkung aufweisen (sog. Totalherbizide), in Kombination mit der Entwicklung von gegenüber dem Totalherbizid resistenten Nutzpflanzen. Gegenüber Totalherbiziden resistente Nutzpflanzen sind bereits verschiedentlich

beschrieben worden. Dabei können mehrere Prinzipien zur Erzeugung einer Resistenz unterschieden werden:

- 5 a) Resistenzerzeugung in einer Pflanze über Mutationsverfahren oder gentechnische Verfahren, indem das als Zielort für das Herbizid dienende Protein deutlich überproduziert wird und indem auf Grund des großen Überschusses des als Zielort für das Herbizid dienende Protein, die von diesem Protein in der Zelle ausgeübte Funktion auch nach Applikation des Herbizides beibehalten wird.
- 10 b) Veränderung der Pflanze dahingehend, dass eine modifizierte Version des als Zielort des Herbizid fungierenden Proteins eingeführt wird und dass das neu eingeführte modifizierte Protein vom Herbizid nicht in seiner Funktion beeinträchtigt wird.
- 15 c) Veränderung der Pflanze dahingehend, dass ein neues Protein/ eine neue RNA eingeführt wird, welche(s) dadurch gekennzeichnet ist, dass die für die herbizide Wirkung der niedermolekularen Substanz verantwortliche chemische Struktur des Proteins oder der Nukleinsäure, wie der RNA oder der DNA, so verändert wird, dass durch die veränderte Struktur keine herbizide Wirkung mehr entfaltet werden kann, oder das Herbizid in der veränderten Pflanze inaktiviert oder modifiziert wird, z.B. abgebaut wird, nicht aufgenommen oder nicht transportiert wird oder in die Vakuole transportiert wird, etc., d.h. die Interaktion des Herbizids mit dem Zielort nicht mehr erfolgen kann.
- 20 d) dass die Funktion des Targets durch eine neue in die Pflanze eingebrachte Nukleinsäure, z.B. ein Gen ersetzt wird, wobei die Nukleinsäure für ein Genprodukt codiert, das geringer oder gar nicht von der herbiziden Substanz in seiner Funktion getrennt wird. So kann z.B. ein sogenannter "alternativer Pathway" geschaffen werden.
- 25 e) dass die Funktion des Targets durch ein anderes in der Pflanze vorhandenes oder in die Pflanze eingebrachtes Gen bzw. dessen Genprodukt übernommen wird.
- 30
- 35 Die vorliegende Erfindung umfasst daher weiterhin die Verwendung von Pflanzen, die durch die T-DNA Insertion getroffene Gene mit den in dem erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen, insbesondere SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33,
- 40

SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 oder die weiteren genannten Sequenzen, z.B. Fragmente und Derivate, z.B. aus anderen Pflanzen, enthalten, zur Entwicklung von neuen Herbiziden. Dem Fachmann sind

5 alternative Verfahren zur Identifizierung von homologen Nukleinsäuren beispielsweise in anderen Pflanzen mit ähnlichen Sequenzen wie unter Verwendung von Transposons, bekannt. Gegenstand dieser Erfindung ist daher auch die Verwendung von alternativen Insertionsmutageneseverfahren zur Insertion von fremder Nukleinsäure in die hierin beschriebenen, erfindungsgemäßen Nukleinsäuresequenzen, insbesondere

10 SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49

15 oder SEQ ID NO: 51 in von diesen Sequenzen aufgrund des genetischen Codes abgeleitete Sequenzen und/oder deren Derivate oder Fragmente z.B. aus anderen Pflanzen.

Ein weiterer Gegenstand der Erfindung sind deshalb wie oben beschrieben Substanzen identifiziert nach den erfindungsgemäßen Verfahren, wobei die Substanz ein Verbindungen vorteilhaft eine niedermolekulare Verbindung ist mit kleiner

20 1000 Molekulargewicht, vorteilhaft kleiner 900 Dalton, bevorzugt kleiner 800, besonders bevorzugt kleiner 700, ganz besonders bevorzugt kleiner 600 Dalton, vorteilhaft mit einem KI-Wert von kleiner 10^{-7} , vorteilhaft kleiner 10^{-8} , bevorzugt kleiner

25 10^{-9} M, vorteilhaft sollte dabei diese Hemmwirkung auf eine spezifische Hemmung der biologischen Aktivität der erfindungsgemäßen Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine zurückzuführen sein, das heißt es sollte keine Hemmung durch diese niedermolekularen Substanzen weiterer nahe verwandter Nukleinsäuren und/oder der durch diese Nukleinsäuren codierten Proteine erfolgen.

30 Weiterhin sollten die niedermolekulare Substanzen vorteilhaft ein Molekulargewicht von größer 50 Dalton, bevorzugt größer 100 Dalton, besonders bevorzugt größer 150 Dalton, ganz besonders bevorzugt größer 200 Dalton haben. Vorteilhaft sollten die niedermolekularen Substanzen weniger als drei Hydroxylgruppen an einem Kohlenstoffatom-enthaltenden Ring haben. Weiterhin sollten auch keine freie(n) Säure- oder

35 Lacton-Gruppe(n) sowie keine Phosphatgruppe und nicht mehr als eine Aminogruppe im Molekül enthalten sein. Auch Basen wie Adenosin im Molekül sind weniger bevorzugt. Bei den Substanzen kann es sich auch vorteilhaft um eine proteinogene Substanz wie einem Antikörper oder um eine Antisense-RNA handelt.

- Eine weitere Ausführungsform der Erfindung sind Substanzen nach den erfindungsgemäßen oben beschriebenen Verfahren identifiziert wurden und wobei es sich bei den Substanzen um einen Antikörper gegen das durch die Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 codierte Protein oder Derivate oder Fragmente davon handelt.
- Die Antikörper können auch mehrere der genannten Sequenzen binden, solange die Bindung spezifisch ist, d.h. nach den oben genannten Verfahren identifiziert oder getestet werden können.
- Diese Substanzen zeichnen sich vorteilhaft durch ihre herbizide Wirkung, die mit den oben beschriebenen Verfahren identifizierbar sind.
- Ein weiterer Erfindungsgegenstand sind Mittel, enthaltend eine herbizid wirksame Menge mindestens einer Substanz identifiziert nach einem der erfindungsgemäßen Verfahren oder eines Antagonisten identifiziert nach einem erfindungsgemäßen Verfahren und mindestens einen inerten flüssigen und/oder festen Trägerstoff sowie gegebenenfalls mindestens einen oberflächenaktiven Stoff.
- Eine weitere Ausführungsform sind Mittel, enthaltend eine das Wachstum regulierende Menge mindestens einer Substanz identifiziert nach den erfindungsgemäßen Verfahren oder eines Antagonisten identifiziert nach einem erfindungsgemäßen Verfahren und mindestens einen inerten flüssigen und/oder festen Trägerstoff sowie gegebenenfalls mindestens einen oberflächenaktiven Stoff.
- Diese erfindungsgemäßen Substanzen oder Mittel mit ihrer herbiziden Wirkung können als Defoliants, Desiccants, Krautabtötungsmittel und insbesondere als Unkrautvernichtungsmittel verwendet werden. Unter Unkraut im weitesten Sinne sind alle Pflanzen zu verstehen, die an Orten aufwachsen, an denen sie unerwünscht sind. Ob die mit Hilfe der erfindungsgemäßen Verfahren gefundenen Substanzen bzw. Wirkstoffe als totale oder selektive Herbizide wirken, hängt unter anderem von der angewandten Menge, ihrer Selektivität und weiteren Faktoren ab. Die Substanzen können beispielsweise gegen folgende Unkräuter verwendet werden:

Dikotyle Unkräuter der Gattungen:

- Sinapis, Lepidium, Galium, Stellaria, Matricaria, Anthemis, Galium, Chenopodium, Urtica, Senecio, Amaranthus, Portulaca, Xanthium, Convolvulus, Ipomoea, Polygonum, Sesbania, Ambrosia, Cirsium, Carduus, Sonchus, Solanum, Rorippa, Rotala, Lindernia, Lamium, Veronica, Abutilon, Emex, Datura, Viola, Galeopsis, Papaver, Centaurea, 5 Trifolium, Ranunculus, Taraxacum.

Monokotyle Unkräuter der Gattungen:

- Echinochloa, Setaria, Panicum, Digitaria, Phleum, Poa, Festuca, Eleusine, Brachiaria, Lolium, Bromus, Avena, Cyperus, Sorghum, Agropyron, Cynodon, Monochoria, 10 Fimbristylis, Sagittaria, Eleocharis, Scirpus, Paspalum, Ischaemum, Sphenoclea, Dactyloctenium, Agrostis, Alopecurus, Apera.

- In Abhängigkeit von der jeweiligen Applikationsmethode können die im erfindungsgemäßen Verfahren identifizierten Substanzen bzw. sie enthaltende Mittel vorteilhaft 15 noch in einer weiteren Zahl von Kulturpflanzen zur Beseitigung unerwünschter Pflanzen eingesetzt werden. In Betracht kommen beispielsweise folgende Kulturen:

- Allium cepa, Ananas comosus, Arachis hypogaea, Asparagus officinalis, Beta vulgaris spec. altissima, Beta vulgaris spec. rapa, Brassica napus var. napus, Brassica napus 20 var. napobrassica, Brassica rapa var. silvestris, Camellia sinensis, Carthamus tinctorius, Carya illinoensis, Citrus limon, Citrus sinensis, Coffea arabica (Coffea canephora, Coffea liberica), Cucumis sativus, Cynodon dactylon, Daucus carota, Elaeis guineensis, Fragaria vesca, Glycine max, Gossypium hirsutum, (Gossypium arboreum, Gossypium herbaceum, Gossypium vitifolium), Helianthus annuus, Hevea brasiliensis, 25 Hordeum vulgare, Humulus lupulus, Ipomoea batatas, Juglans regia, Lens culinaris, Linum usitatissimum, Lycopersicon lycopersicum, Malus spec., Manihot esculenta, Medicago sativa, Musa spec., Nicotiana tabacum (N. rustica), Olea europaea, Oryza sativa, Phaseolus lunatus, Phaseolus vulgaris, Picea abies, Pinus spec., Pisum sativum, Prunus avium, Prunus persica, Pyrus communis, Ribes sylestre, Ricinus 30 communis, Saccharum officinarum, Secale cereale, Solanum tuberosum, Sorghum bicolor (s. vulgare), Theobroma cacao, Trifolium pratense, Triticum aestivum, Triticum durum, Vicia faba, Vitis vinifera, Zea mays.

- Die nach dem erfindungsgemäßen Verfahren gefundenen Substanzen können 35 vorteilhaft auch in Kulturen verwandt werden, die durch Züchtung einschließlich gentechnischer Methoden gegen die Wirkung von Herbiziden tolerant sind.

- Die erfindungsgemäßen Substanzen bzw. die sie enthaltenden herbiziden Mittel können beispielsweise in Form von direkt versprühbaren wässrigen Lösungen, 40 Pulvern, Suspensionen, auch hochprozentigen wässrigen, öligen oder sonstigen

Suspensionen oder Dispersionen, Emulsionen, Öldispersionen, Pasten, Stäubemitteln, Streumitteln oder Granulaten durch Versprühen, Vernebeln, Verstäuben, Verstreuen oder Gießen angewendet werden. Die Anwendungsformen richten sich nach den Verwendungszwecken; sie sollten in jedem Fall möglichst die feinste Verteilung der erfindungsgemäßen Wirkstoffe gewährleisten.

Als inerte flüssige und/oder feste Trägerstoffe kommen flüssige Zusatzstoffe wie Mineralölfractionen von mittlerem bis hohem Siedepunkt, wie Kerosin oder Dieselöl, ferner Kohlenteeröle sowie Öle pflanzlichen oder tierischen Ursprungs, aliphatische, cyclische und aromatische Kohlenwasserstoffe, z.B. Paraffin, Tetrahydronaphthalin, alkylierte Naphthaline oder deren Derivate, alkylierte Benzole oder deren Derivate, Alkohole wie Methanol, Ethanol, Propanol, Butanol, Cyclohexanol, Ketone wie Cyclohexanon oder stark polare Lösungsmittel, z.B. Amine wie N-Methylpyrrolidon oder Wasser in Frage.

Weitere vorteilhafte Anwendungsformen der erfindungsgemäßen Substanzen und/oder Mittel sind wässrige Anwendungsformen wie Emulsionskonzentrate, Suspensionen, Pasten, netzbaren Pulvern oder wasserdispersierbaren Granulaten, die beispielsweise durch Zusatz von Wasser bereitet werden können. Zur Herstellung von Emulsionen, Pasten oder Öldispersionen können die Substanzen und/oder Mittel die sog. Substrate als solche oder in einem Öl oder Lösungsmittel gelöst, mittels Netz-, Haft-, Dispersier- oder Emulgiermittel in Wasser homogenisiert werden. Es können aber auch aus wirksamer Substanz, Netz-, Haft-, Dispersier- oder Emulgiermittel und eventuell Lösungsmittel oder Öl bestehende Konzentrate hergestellt werden, die zur Verdünnung mit Wasser geeignet sind.

Als oberflächenaktive Stoffe kommen die Alkali-, Erdalkali-, Ammoniumsalze von aromatischen Sulfonsäuren, z.B. Lignin-, Phenol-, Naphthalin- und Dibutylnaphthalin-sulfonsäure, sowie von Fettsäuren, Alkyl- und Alkylarylsulfonaten, Alkyl-, Laurylether- und Fettalkoholsulfaten, sowie Salze sulfatierter Hexa-, Hepta- und Octadecanolen sowie von Fettalkoholglykoether, Kondensationsprodukte von sulfoniertem Naphthalin und seiner Derivate mit Formaldehyd, Kondensationsprodukte des Naphthalins bzw. der Naphthalinsulfonsäuren mit Phenol und Formaldehyd, Polyoxyethylenoctylphenol-ether, ethoxyliertes Isooctyl-, Octyl- oder Nonylphenol, Alkylphenyl-, Tributylphenyl-polyglykoether, Alkylarylpolyetheralkohole, Isotridecylalkohol, Fettalkoholethylenoxid-Kondensate, ethoxyliertes Rizinusöl, Polyoxyethylenalkylether oder Polyoxypropylen-alkylether, Laurylalkoholpolyglykoetheracetat, Sorbitester, Lignin-Sulfitaugen oder Methylcellulose in Betracht.

Pulver-, Streu- und Stäubemittel können als feste Trägerstoffe teilweise durch Mischen oder gemeinsames Vermahlen der wirksamen Substanzen mit einem festen Trägerstoff hergestellt werden.

- 5 Granulate, z.B. Umhüllungs-, Imprägnierungs- und Homogengranulate können durch Bindung der Wirkstoffe an feste Trägerstoffe hergestellt werden. Feste Trägerstoffe sind beispielsweise Mineralerden wie Kieselsäuren, Kieselgele, Silikate, Talkum, Kaolin, Kalkstein, Kalk, Kreide, Bolus, Löß, Ton, Dolomit, Diatomeenerde, Calcium- und Magnesiumsulfat, Magnesiumoxid, gemahlene Kunststoffe, Düngemittel, wie
- 10 Ammoniumsulfat, Ammoniumphosphat, Ammoniumnitrat, Harnstoffe und pflanzliche Produkte wie Getreidemehl, Baumrinden-, Holz- und Nussschalenmehl, Cellulosepulver oder andere feste Trägerstoffe.

- 15 Die Konzentrationen der erfindungsgemäßen Substanzen und/oder Mittel in den anwendungsfertigen Zubereitungen können in weiten Bereichen variiert werden. Die Formulierungen enthalten im allgemeinen 0,001 bis 98 Gew.-%, vorzugsweise 0,01 bis 95 Gew.-%, mindestens eines Wirkstoffs. Die Wirkstoffe werden dabei in einer Reinheit von 90 % bis 100 %, vorzugsweise 95 % bis 100 % (nach NMR-Spektrum) eingesetzt.

- 20 Die Applikation der herbiziden Mittel bzw. der Substanzen kann im Vorauf- oder im Nachaufverfahren erfolgen. Sind die Wirkstoffe für gewisse Kulturpflanzen weniger verträglich, so können Ausbringungstechniken angewandt werden, bei welchen die herbiziden Mittel oder Substanzen mit Hilfe der Spritzgeräte so gespritzt werden, dass die Blätter der empfindlichen Kulturpflanzen nach Möglichkeit nicht getroffen werden,
- 25 während die Wirkstoffe auf die Blätter darunter wachsender unerwünschter Pflanzen oder die unbedeckte Bodenfläche gelangen (post-directed, lay-by).

- Zur Verbreiterung des Wirkungsspektrums und zur Erzielung synergistischer Effekte können die erfindungsgemäßen Substanzen und/oder Mittel mit zahlreichen Vertretern
- 30 anderer herbizider oder wachstumsregulierender Wirkstoffgruppen gemischt und gemeinsam ausgebracht werden. Beispielsweise kommen als Mischungspartner 1,2,4-Thiadiazole, 1,3,4-Thiadiazole, Amide, Aminophosphorsäure und deren Derivate, Aminotriazole, Anilide, (Het)-Aryloxyalkansäure und deren Derivate, Benzoesäure und deren Derivate, Benzothiadiazinone, 2-Aroyl-1,3-cyclohexandione, Hetaryl-Aryl-Ketone,
- 35 Benzylisoxazolidinone, Meta-CF₃-phenylderivate, Carbamate, Chinolinsäure und deren Derivate, Chloracetanilide, Cyclohexan-1,3-dionderivate, Diazine, Dichlorpropionsäure und deren Derivate, Dihydrobenzofurane, Dihydrofuran-3-one, Dinitroaniline, Dinitrophenole, Diphenylether, Dipyridyle, Halogencarbonsäuren und deren Derivate, Harnstoffe, 3-Phenyluracile, Imidazole, Imidazolinone, N-Phenyl-3,4,5,6-tetrahydrophthalimide, Oxadiazole, Oxirane, Phenole, Aryloxy- oder Heteroaryloxyphenoxypropion-
- 40

säureester, Phenylacessigsäure und deren Derivate, Phenylpropionsäure und deren Derivate, Pyrazole, Phenylpyrazole, Pyridazine, Pyridincarbonsäure und deren Derivate, Pyrimidylether, Sulfonamide, Sulfonylhamstoffe, Triazine, Triazinone, Triazolinone, Triazolcarboxamide, Uracile in Betracht.

5

Außerdem kann es von Nutzen sein, die erfindungsgemäßen Substanzen und/oder Mittel allein oder in Kombination mit anderen Herbiziden auch noch mit weiteren Pflanzenschutzmitteln gemischt, gemeinsam auszubringen, beispielsweise mit Mitteln zur Bekämpfung von Schädlingen oder phytopathogenen Pilzen bzw. Bakterien. Von Interesse ist ferner die Mischbarkeit mit Mineralsalzlösungen, welche zur Behebung von Ernährungs- und Spurenelementmängeln eingesetzt werden. Es können auch nichtphytotoxische Öle und Ölkonzentrate zugesetzt werden.

10

Die Aufwandmengen an Wirkstoff (= Substanzen und/oder Mittel) betragen je nach Bekämpfungsziel, Jahreszeit, Zielpflanzen und Wachstumsstadium 0,001 bis 3,0, vorzugsweise 0,01 bis 1,0 kg/ha aktive Substanz.

15

Ein weiterer erfindungsgemäßer Gegenstand ist die Verwendung einer Substanz identifiziert nach einem der erfindungsgemäßen Verfahren oder Mittel enthaltend diese Substanzen als Herbizid oder zur Wachstumsregulierung von Pflanzen.

20

Außerdem betrifft die Erfindung ein Kit, umfassend das erfindungsgemäße Nukleinsäurekonstrukt, der erfindungsgemäßen Substanzen, z.B. den erfindungsgemäßen Antikörper, das erfindungsgemäße Antisense-Nukleinsäuremolekül und/oder einen Antagonisten und/oder eine herbizide Substanz identifiziert gemäß den erfindungsgemäßen Verfahren sowie die unten beschriebene Zusammensetzung.

25

Ein weiterer erfindungsgemäßer Gegenstand ist eine Zusammensetzung, enthaltend die erfindungsgemäße Substanz, den erfindungsgemäßen Antikörper, das erfindungsgemäße Antisense-Nukleinsäurekonstrukt und/oder einen erfindungsgemäßen Antagonisten und/oder eine erfindungsgemäße Substanz identifiziert nach einem erfindungsgemäßen Verfahren.

30

Die Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als beschränkend aufgefasst werden sollten.

35

Beispiele

a) Molekularbiologische Methoden

40

Molekularbiologische Methoden wie sie hier eingesetzt werden entsprechen dem Stand der Technik und sind an verschiedenen Stellen, wie bspw. Sambrook et al., Molecular Cloning, eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), Reiter et al., Methods in Arabidopsis Research, World Scientific Press (1992), Schultz et al., Plant Molecular Biology Manual, Kluwer Academic Publishers (1998) und Martinez-Zapater und Salinas, Methods in Molecular Biology, Vol. 82: Arabidopsis Protocols eds., Humana Press Inc., Totowa, NJ. Diese Referenzen beschreiben die gängigen Standardmethoden für die Produktion, Identifizierung und Klonierung von durch T-DNA-Insertionen hervorgerufenen Mutanten. Zusätzlich wurde für die Identifizierung von Insertionsstellen auf eine weitere gängige Methode, wie sie u.a. von Spertini et al., Biotechniques 27: 308-314 (1999) beschrieben wurde, zurückgegriffen. Die Sequenzierungen erfolgten durch die Firma DNA LandMarks Inc., Quebec, Kanada.

15

b) Material

Die verwendeten Chemikalien wurden, wenn im Text nicht gesondert angegeben, in p.A.-Qualität von den Firmen Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deideshofen) bezogen. Lösungen wurden unter Verwendung von reinem pyrogenfreiem Wasser, das mit einer Ionenaustauschanlage der Firma (TKA, Niederelbert) hergestellt wurde. Restriktionsnukleasen, DNA-modifizierende Enzyme und molekularbiologische Kits und Oligonukleotide wurden bezogen von den Firmen Amersham Pharmacia (Freiburg), Biometra (Göttingen), Dynal (Hamburg), Gibco-BRL (Gaithersburg, MD., USA), Invitrogen (Groningen, Niederlande), MBI Fermentas (St. Leon Rot), New England Biolabs (Schwalbach, Taunus), Novagen (Madison, Wisconsin, USA), Qiagen (Hilden), Roche Diagnostics (Mannheim), Stratagene (Amsterdam, Niederlande), TIB-Molbiol (Berlin). Falls nicht anders beschrieben wurden die Produkte nach den Angaben der Hersteller eingesetzt.

30

Beispiel 1: Herstellung einer KO-Population und Identifizierung von Linien, die für lethale Mutation segregieren

Ausgehend von dem Grundgerüst der pPZP-Vektoren [Hajukiewicz, P. et al., (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol. Biol. 25, 989-994], wurde ein modifizierter binärer Vektor konstruiert, welcher das Kanamycin-Resistenzgen für die Selektion in Bakterien enthielt. Zwischen der linken und der rechten T-DNA-Border befand sich nur eine Selektionskassette, bestehend aus dem Resistenz Gen für die Clearfieldresistenz (Imidazolinon-

40

bzw. AHAS-Resistenz) unter der Kontrolle des konstitutiven Promotors *mas1* (Velten et al., 1984, EMBO J. 3, 2723-2730; Mengiste, Amedeo and Paszkowski, 1997, Plant J., 12, 945-948.). Alternativ können anderen Resistenzgene wie die Herbizidresistenzgene wie die Phosphinotricin- (= bar-Resistenz), die Methioninsulfoximin-, die Sulfonyl-harnstoff- (= ilv-Resistenz, ind *S. cerevisiae* ilv2) oder die Phenoxypyphenoxypyherbizid-resistenzgene (= ACCase-Resistenz) oder Antibiotikaresistenzgene verwendet werden. Auch sind dem Fachmann anderen konstitutive Promotoren bekannt, die anstelle des verwendeten *mas1*-Promotors verwendet werden können wie der 34S-, der 35S- oder der Ubiquitin-Promotor aus *Petersilie*. Dem Fachmann sind die verschiedenen für die Transformation von *Arabidopsis* mittels *Agrobacterium* verwendbaren Vektoren bekannt. Eine detaillierte Beschreibung der einsetzbaren Vektoren und *Agrobacterien*-stämme findet man in Hellens et al., (Trends in Plant Science, 2000, Vol 5, 446-451). Die Plasmide wurden mittels eines Hitze-Schock Protokolls in *Agrobacterien* hier in den *Agrobacterium tumefaciens* Stamms GV3101pMP90 (Koncz and Schell, 1986 Mol. Gen. Genet. 204:383-396) transformiert. Transformierte Kolonien der Bakterien wurden auf YEP-Medium, welche die entsprechenden Antibiotika enthielten, für 2 Tage bei 28°C angezogen. Diese *Agrobacterien* wurden dann für die Transformationen einer großen Zahl an *Arabidopsis*-Pflanzen des Ökotyps C24 (Nottingham *Arabidopsis* Stock Centre, UK ; NASC Stock N906) eingesetzt, wobei nach einer modifizierten Version der "in-planta" Transformationsmethode (Bechtold, N., Ellis, J., Pelletier, G. 1993. In planta *Agrobacterium* mediated gene transfer by infiltration of *Arabidopsis thaliana* plants, C.R. Acad. Sci. Paris. 316:1194-1199; Clough, JC and Bent, AF. 1998 Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*, Plant J. 16:735-743) verfahren wurde. Transformierte Pflanzen wurden mittels des Selektionsagens, gegen welches das auf der T-DNA codierte Resistenzgen eine Resistenz vermittelt, selektiert.

Von diesen transformierten Pflanzen wurden ca. 100 bis 200 Samen (T2) auf Agarplatten mit Selektionsagens ausgelegt. Diese Platten wurden für 2 Tage bei 4°C stratifiziert und für ca. 7 bis 10 Tage bei 20°C und Dauerlicht inkubiert. Anschließend wurde die Zahl an gegenüber dem Selektionsagens resistenten und sensitiven Keimlingen bestimmt. Zudem wurden gegebenenfalls die Zahl an nicht pigmentierten Pflanzen (Albinos) bestimmt. Diese ließen sich auf Grund ihrer Farbe eindeutig von den sensitiven Keimlingen unterscheiden. Es wurden nur solche Linien weiter untersucht, die offensichtlich für einen Insertionsort segregierten, das heißt in denen ungefähr ein Drittel bis zu einem Viertel der Pflanzen eine Sensitivität gegenüber der Selektion zeigten und in denen eine sehr enge Kopplung, das heißt, dass eine Cosegregation zwischen der die Resistenz vermittelnde T-DNA und der den Phänotyp erzeugenden Mutation gefunden wurde. Eine solche sehr enge Kopplung zwischen der T-DNA und der Mutation lag vor, wenn ein Zahlenverhältnis von 2:1 zwischen

resistenten und sensitiven Keimlingen gefunden wurde. Dies von einer normalen 3:1 Spaltung für einen Insertionsort abweichende Zahlenverhältnis kommt nur vor, wenn die homozygot resistenten Pflanzen quantitativ fehlen, weil sie entweder schon als Embryo sterben bzw. sich nicht entwickeln oder aber weil sie einen Albinophänotyp ausprägen. Demnach ist es sehr wahrscheinlich, dass die Insertion der T-DNA an dem jeweiligen Ort im Genom die Ursache für die embryonalelethale bzw. die Albinomutation darstellt. Entsprechend kann das essentielle Gen durch Bestimmung des Insertionsort und des dort vorhandenen Gens identifiziert werden.

10 Beispiel 2: Molekulare Analyse von Linien mit embryo- oder albinolethalem Phänotyp

Von den ausgewählten Linien, die für eine albinolethale oder embryoletale Mutation segregierten und für die eine Cosegregation zwischen T-DNA und Mutation bestimmt wurde, wurde aus ca. 50 mg Blattmaterial genomische DNA mittels Standardverfahren (entweder Säulen der Firma Qiagen, Hilden, Germany oder Phytopure Kit der Firma Amersham Pharmacia, Freiburg, Germany) isoliert. Die Amplifikation der Insertionsstelle der T-DNA erfolgte durch eine modifizierte Version der Adaptor-PCR-Methode wie sie von Spertini D, Béliveau C. and Bellemare, 1999, Biotechniques, 27, 308-314 publiziert worden ist. Etwa jeweils 50 bis 100-ng der genomischen DNA wurden parallel mit den Restriktionsenzymen MunI, BglII, BspI (= Bsp119I), PspI (= Psp1406I) und SspI verdaut und mit einem Adaptor ligiert, der aus den annealten Oligos 5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGGGCAGGT-3' und 5'-NN(2-4)ACCTGCCCAA-3' bestand, wobei 5'-NN₍₂₋₄₎ den zum jeweiligen Enzym passenden Überhang repräsentierten. Ein µl dieser mit Adaptoren versehenen genomischen DNA wurde für eine Amplifikation der zu den T-DNAs flankierend liegenden Sequenzen unter Verwendung eines adaptorspezifischen (5'-GGATCCTAATACGACTCACTATAGGGC-3') und jeweils eines genspezifischen Primers für jede Border eingesetzt. Dem Fachmann ist bekannt, wie man genspezifische Primer für die T-DNA, welche zur Pflanzentransformation verwendet wurde, designt und synthetisiert. Die PCR wurde unter Standardbedingungen für 7 Zyklen bei einer Annealingtemperatur von 72°C und für 32 Zyklen bei einer Annealingtemperatur von 65°C in 25 µl Reaktionsvolumen durchgeführt. Das erhaltene Amplifikat wurde 1:50 in H₂O verdünnt und ein µl dieser Verdünnung für eine zweite Amplifikationsrunde (5 Zyklen bei einer Annealing-Temperatur von 67°C und 28 Zyklen bei einer Annealingtemperatur von 60°C) eingesetzt. Hierzu wurden "eingerückte" daher auf dem PCR-Produkt weiter innenliegende Primer eingesetzt, wodurch die Spezifität und Selektivität der Amplifikation erhöht wurde. Von dem in 50 µl Reaktionsvolumen erhaltenen Amplifikaten wurde ein Aliquot einer gelelektrophoretischen Analyse unterzogen. Es wurden jeweils ein oder mehrere spezifische PCR-Produkte für die linke und/oder die rechte T-DNA erhalten. Die Produkte wurden mittels Standard-

methoden (Firmingen, Hilden) gereinigt und mit Hilfe weit T-DNA-spezifischer Primer sequenziert. Durch einen Blast-Vergleich (BLASTN, Altschul, et al., 1990, J Mol. Biol. 215:403-410) der isolierten Sequenz mit der publizierten Genomsequenzen von Arabidopsis (The Arabidopsis Genome Initiative, 2000, Nature, 408:796-815) wurde
5 jeweils die Insertionsstelle der T-DNA im Genom bestimmt. Da diese Sequenzen in annotierten Form in verschiedenen Datenbanken, die dem Fachmann bekannt sind, zur Verfügung stehen, konnten auch die jeweils inaktivierten ORFs bestimmt werden. Die erfolgreiche Identifizierung eines inaktivierten ORFS wurde durch eine PCR-Reaktion mit einem für die abgeleitete flankierenden Sequenz und einem für die T-DNA
10 spezifischen Primer verifiziert. Der Erhalt des PCR-Produkts der erwarteten Größe spezifisch für diese Linie bestätigte die erfolgreiche Identifizierung der Insertionsstelle der T-DNA.

15 Beispiel 3: Identifizierung und Analyse der Linie 303317 die für eine lethale Mutation segregiert

Die Linie 303317 wurde wie oben beschrieben (Beispiele 1 und 2) als Linie, die für eine keimlingslethale Mutation segregiert, identifiziert. Die genaue Auszählung der Spaltung ergab, dass 25 % der Nachkommen den Albinophänotyp, 25 % der Nachkommen
20 Sensitivität gegenüber der Selektion und 50 % der Nachkommen Resistenz gegenüber der Selektion zeigten. Dieses Spaltungsverhältnis wird erwartet, wenn ausschließlich die homozygot resistenten Keimlinge den Phänotyp zeigen, daher die T-DNA-Insertion mit der lethalen Mutation sehr eng gekoppelt ist. Die Kopplung wurde weiterhin in einer Cosegregationsanalyse überprüft. Dazu wurden von 40 wildtypischen, resistenten
25 Pflanzen der Linie 303317 die Nachkommen analysiert. In allen Fällen konnten wiederum Albinos in der Nachkommenschaft gefunden werden. Dieser Umstand lässt den Schluss zu, dass die Resistenz-vermittelnde T-DNA-Insertion und die Mutation immer zusammen vererbt werden und damit (mit sehr hoher Wahrscheinlichkeit) übereinstimmen. Die molekularbiologische Analyse wurde wie für Beispiel 1 beschrieben durchgeführt. Für die Linie 303317 wurde für das Enzym MunI ein 1400 bp-großes
30 Fragment für die linke T-DNA-Border identifiziert. Der Erhalt des PCR-Produkts der vorhergesagten Größe spezifisch für diese Linie bestätigte die erfolgreiche Identifizierung der Insertionsstelle der T-DNA. Die Blast-Analyse der isolierten Sequenz (BLASTN, Altschul et al., 1990) J Mol. Biol. 215:403-410) demonstrierte die Insertion
35 der T-DNA in Position 6628 des BAC-Klons ATF2809 mit der Accession Number AL137080. Nach der Annotation dieses Bereichs ist die Integration in einen ORF (F2809.40, SEQ ID NO: 1) erfolgt, der Ähnlichkeit zum "Translation Releasing Factor" RF-2 aus Synechocystis sp. (PIR:S76448) hat. Weiterhin besitzt das Protein (SEQ ID NO: 2) eine araC-Familiensignatur. Die erfolgreiche Identifizierung der Insertionsstelle

und des inaktivierten ORFs wurde durch eine PCR-Reaktion mit einem für die abgeleitete flankierende Sequenz und einem für die T-DNA-spezifischen Primer verifiziert.

- 5 Beispiel 4: Identifizierung und Analyse der Linien 304149, 120701, 126548, 127023, 127235, 218031, 171042, KO-T3-02-33338-3, KO-T3-02-33885-2 und KO-T3-02-35172-2, die für eine lethale Mutation segregieren

10 Analog den vorgenannten Beispielen 1 bis 4 wurden die Klone 304149, 120701, 126548, 127023, 127235, 218031, 171042, KO-T3-02-33338-3, KO-T3-02-33885-2 und KO-T3-02-35172-2, als Linien identifiziert, die für embryo- oder keimlingslethale Mutationen segregieren. Die Aufspaltung erfolgte bei allen Linien wie unter Beispiel 3 bzw. analog zu Beispiel 3 für embryoletale Mutationen beschrieben. Die embryoletale Mutation führt jedoch dazu, dass die für die Mutation homozygoten Pflanzen bereits im Embryostadium ihre Entwicklung abbrechen, somit gar nicht keimen. Entsprechend
15 verschiebt sich das Zahlenverhältnis nach einem Drittel sensitive und zwei Drittel resistente Pflanzen gegenüber der Selektion. Die molekularbiologischen Arbeiten bzw. Analysen wurden wie unter Beispiel 1 bis 3 beschrieben durchgeführt.

20 Die Linie 304149 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 304149 wurde für das Enzym MunI ein 750 bp-großes Fragment, für die Enzyme Psp1406I/Bsp119I ein 300 bp-großes Fragment und für das Enzym SpeI ein 950 bp großes Fragment jeweils für die linke T-DNA-Border identifiziert. Für die rechte T-DNA-Border wurde mit dem Enzym SpeI ein 300 bp großes Fragment identifiziert. Die Sequenzierung dieser Fragmente
25 offenbarte den gleichen Insertionsort. Die T-DNA ist auf Chromosom 5 in Position 35398 des P1-Klons MSH12, Accession AB006704 inseriert. Durch die Insertion 110 bp stromaufwärts vom Startcodon des ORFs MSH12.9 wird höchstwahrscheinlich die Transkription unterbunden bzw. die Transkriptstabilität verringert und damit die Funktionalität des ORFs verringert oder vollständig zerstört. Dieser ORF MSH12.9
30 codiert für ein Kobalamin-Synthese-Protein.

Die Linie 120701 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 120701 wurde für das Enzym BglII ein 500 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die T-DNA
35 ist auf Chromosom 4 in Position 55170 des Bac-Klons ATT25K17, Accession AL049171 inseriert. Durch die Insertion innerhalb des codierenden Bereichs wird der ORF T25K17.110 unterbrochen und damit inaktiviert. Dieser ORF T25K17.110 codiert für eine Arginyl-tRNA-Synthetase. Dieser ORF beinhaltet den EST: gb:AA404880, T76307.

Die Linie 126548 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 126548 wurde für die Enzyme Psp1406I/Bsp119I ein 1000 bp-großes Fragment für die linke T-DNA-Border identifiziert. Für die rechte T-DNA-Border wurden mit den Enzymen Psp1406I/Bsp119I ein
5 900 bp-großes Fragment und mit dem Enzym BglII ein 300 bp-großes Fragment identifiziert. Die Sequenzierung aller PCR-Produkte demonstrierte die Insertion der T-DNA am gleichen genomischen Ort. Die T-DNA ist auf Chromsom 4 in Position 36872 des Bac-Klons ATF17A8, Acession AL049482 inseriert. Durch die Insertion innerhalb des codierenden Bereichs wird der ORF F17A8.80 unterbrochen und damit
10 inaktiviert. Dieser ORF F17A8.80 codiert für putatives Protein mit Ähnlichkeit zu einer RNA-Helicase aus Maus (*Mus musculus*), PIR2:184741.

Die Linie 127023 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 127023 wurde für das Enzym
15 BglII ein 350 bp-großes Fragment und für die Enzyme Psp1406I/Bsp119I ein 900 bp-großes Fragment jeweils für die linke T-DNA-Border identifiziert. Beide Fragmente identifizierten nach der Sequenzierung den identischen Insertionsort. Die T-DNA ist auf Chromsom 4 in Position 61403 des BAC-Klons ATT19P19, Acession AL022605 inseriert. Durch diese Insertion wird der ORF AT4g39780 unterbrochen und damit
20 inaktiviert. Dieser ORF AT4g39780 codiert für ein putatives Protein mit Ähnlichkeit zum dem die AP2-Domäne enthaltenden Protein RAP 2.4 aus *Arabidopsis thaliana*. Zudem beinhaltet dieser ORF die ESTs gb:T46584 und AA394543.

Die Linie 127235 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 127235 wurde für das Enzym
25 MunI ein 1600 bp-großes Fragment für die linke T-DNA-Border identifiziert. Für die rechte T-DNA-Border wurde mit dem Enzym BglII ein 600 bp-großes Fragment identifiziert. Beide Fragmente identifizierten nach der Sequenzierung den identischen Insertionsort. Die T-DNA ist auf Chromsom 1 in Position 10776 des BAC-Klons F9K20, Accession AC005679 inseriert. Durch diese Insertion wird der ORF F9K20.4 unter-
30 brochen und damit inaktiviert. Dieser ORF F9K20.4 codiert für ein putatives Protein mit Ähnlichkeit zu dem gil1786244 hypothetischen 24.9 kD Protein in der *surA*-*hepA* intergenischen Region *yab0* des *Escherichia coli* Genoms gb:AE000116, und zu dem hypothetischen Protein der YABO Familie PFI00849. Weiterhin besitzt das durch den
35 ORF F9K20.4 codierte Protein eine konservierte Pseudouridylat-Synthase-Domäne, welche an der Modifizierung von Uracil in RNA-Molekülen beteiligt ist. Demnach zeigt der ORF F9K20.4 im blastp-Vergleich unter Standardbedingungen signifikante Homologie zu verschiedenen Pseudouridylat-Synthasen.

Die Linie 218031 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 218031 wurde für das Enzym BglII ein 400 bp-großes Fragment für die linke T-DNA-Border identifiziert und dieses anschließend sequenziert. Die T-DNA ist auf Chromsom 2 in Position 11909 des Klons F3G5 mit der Accession AC005896 inseriert. Durch die Insertion im codierenden Bereich wird der ORF At2g37250 inaktiviert. Dieser ORF codiert für eine putative Adenylatkinase.

Die Linie 171042 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 171042 wurde für die Enzyme Psp1406I/Bsp119I ein 1600 bp-großes Fragment für die linke T-DNA-Border identifiziert und dieses anschließend sequenziert. Die T-DNA ist auf Chromsom 3 in Position 97005 des Bac-Klons T29H11 mit der Accession AL049659 inseriert. Durch die Insertion im codierenden Bereich wird der ORF T29H11_270 inaktiviert. Dieser ORF T29H11_270 codiert für ein putatives Protein mit Ähnlichkeit zu dem pol Polyprotein des equine infectious anemia virus (PIR:GNLJEV).

Die Linie KO-T3-02-33338-3 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-33338-3 wurde für das Enzym MunI ein 624 bp-großes Fragment für die linke T-DNA-Border identifiziert und dieses anschließend sequenziert. Die T-DNA ist auf Chromsom 5 in Position 39500 des P1-Klons MJE7 mit der Accession AB020745 inseriert. Durch die Insertion 64 Basenpaare stromabwärts vom Stop-Codon des ORFs MEJ7.11 wird vermutlich das Transkript dieses ORFs verändert und damit die Transkriptstabilität reduziert. Demnach ist von einer reduzierten oder gänzlich blockierten Genfunktion für diesen ORF auszugehen. Der ORF MEJ7.11 codiert für ein unbekanntes Protein.

Die Linie KO-T3-02-33885-2 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-33885-2 wurde für die Enzyme Psp1406I/Bsp119I ein 450 bp-großes Fragment für die linke T-DNA-Border identifiziert. Für die rechte T-DNA-Border wurde mit den Enzymen Psp1406I/Bsp119I ein 650 bp-großes Fragment identifiziert. Beide Fragmente identifizierten nach der Sequenzierung den identischen Insertionsort. Die T-DNA ist auf Chromsom 1 in Position 76356 des Bac-Klons F14G9 mit der Accession AC069159 inseriert. Durch die Insertion im codierenden Bereich des ORFs F14G9.26 wird dieser ORF in dieser Linie inaktiviert. Der ORF F14G9.26 codiert für ein unbekanntes Protein.

Die Linie KO-T3-02-35172-2 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-35172-2

wurde für das Enzym MunI ein 700 bp-großes Fragment für die linke T-DNA-Border identifiziert und dieses anschließend sequenziert. Die T-DNA ist auf Chromosom 5 in Position 24422 des P1-Klons MAB16 mit der Accession AB018112 inseriert. Durch diese Insertion 87bp stromaufwärts des ORFs MAB16.6 wird die Transkription dieses ORFs höchstwahrscheinlich blockiert und damit das Gen ausgeschaltet. Der ORF MAB16.6 codiert für ein Protein, welches nur Homologie zu anderen unbekannten Proteinen zeigt.

Beispiel 5: Identifizierung und Analyse der Linien 305861, 303814, KO-T3-02-132241, KO-T3-02-15114-2, KO-T3-02-18601-1 und 304143, die für albinolethale Mutationen segregieren

Analog den vorgenannten Beispielen 1 bis 4 wurden die Klone 305861, 303814, KO-T3-02-132241, KO-T3-02-15114-2, KO-T3-02-18601-1 und 304143 als Linien identifiziert, die für albinolethale Mutationen segregieren. Die Aufspaltung erfolgte bei allen Linien wie unter Beispiel 3 beschrieben. Die molekular-biologischen Arbeiten bzw. Analysen wurden wie unter Beispiel 1 bis 3 beschrieben durchgeführt.

Die Linie 305861 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 305861 wurde für die Enzymkombination Bgl II ein ca. 1300 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 16326 des BACs T7B11, Accession AC007138 auf Chromosom 4. Durch die Insertion in das offene Leseraster wird der ORF T7B11.6 unterbrochen und inaktiviert. Dieser ORF kodiert für ein „preprotein translocase secA precursor“ Protein, es ist daher ein chloroplastidäres SecA-Protein, das für die Proteintransport über die Thylakoidmembran verantwortlich ist. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

Die Linie 303814 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 303814 wurde für die Enzymkombination Mun I ein ca. 1300 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 2027 des BACs F2G19, Accession AC083835 auf Chromosom 1. Durch die Insertion in das offene Leseraster wird der ORF F2G19.1 unterbrochen und inaktiviert. Dieser ORF kodiert für ein Protein mit signifikanter Homologie zum DCL Protein aus Tomate, PIR:S71749. Weiterhin besitzt das Protein eine sogenannte die HMG-Signatur, der high mobility group Proteine, welche an DNA

binden können. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- 5 Die Linie KO-T3-02-13224-1 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-13224-1 wurde für die Enzymkombination Bgl II ein ca. 500 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 55170 des
- 10 BACs T25K17, Accession AL049171 auf Chromosom 4. Durch die Insertion in das offene Leseraster wird der ORF T25K17.110 unterbrochen und inaktiviert. Dieser ORF kodiert für eine Arginin-tRNA-Ligase. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.
- 15 Die Linie KO-T3-02-15114-2 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-15114-2 wurde für die Enzymkombination Mun I ein ca. 350 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die
- 20 Insertion der T-DNA in dieser Linie an Basenpaarposition 6984 des BACs T5N23, Accession AL138650 auf Chromosom 3. Durch die Insertion in das offene Leseraster wird der ORF T5N23.20 unterbrochen und inaktiviert. Dieser ORF, kodiert für eine plastidäre Gluthathion-Reduktase. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem
- 25 ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.
- Die Linie KO-T3-02-18601-1 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-18601-1 wurde für die Enzymkombination Bgl II ein ca. 600 bp-großes Fragment für die rechte
- 30 T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 4026 des BACs F22O13, Accession AC003981 auf Chromosom 1. Durch die Insertion in das offene Leseraster wird der ORF F22O13.2 unterbrochen und inaktiviert. Dieser ORF kodiert einen
- 35 Transkriptionsinitiationsfaktor Sigma Homolog, daher einem pflanzlichen Homologen zum Sigma-Untereinheit der bakteriellen RNA-Polymerase. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

Die Linie 30414 segregiert für eine albinolethale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie 304143 wurde für das Enzym Bgl II ein ca. 950 bp-großes Fragment für die rechte T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 79156 des BACs F9O13 map mi398, Accession AC006248 auf Chromosom 2. Durch die Insertion in den Promotor, daher ca. 450bp stromaufwärts vom Startcodon, wird vermutlich die Transkription des ORFs At2g15680 verhindert und damit die Genfunktion ausgeschaltet. Der ORF At2g15680 kodiert für ein putatives Calmodulin-ähnliches Protein. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

Beispiel 6: Identifizierung und Analyse der Linien KO-T3-02-403222-2, KO-T3-02-40309-1, KO-T3-02-40309-2, KO-T4-02-00666-4, KO-T4-02-00666-5, KO-T3-02-41568-2, KO-T3-02-42903-1, KO-T3-02-41395-1 und KO-T3-02-44634-4, die für embryoletale Mutationen segregieren

Analog den vorgenannten Beispielen 1 bis 4 wurden die Klone KO-T3-02-403222-2, KO-T3-02-40309-1, KO-T3-02-40309-2, KO-T4-02-00666-4, KO-T4-02-00666-5, KO-T3-02-41568-2, KO-T3-02-42903-1, KO-T3-02-41395-1 und KO-T3-02-44634-4 als Linien identifiziert, die für embryoletale Mutationen segregieren.

Diesen Linien spalten analog zu dem für keimlingslethale Linien beschriebenen Beispiel 3 auf. Die embryoletale Mutation führt jedoch dazu, dass die für die Mutation homozygoten Pflanzen bereits im Embryostadium ihre Entwicklung abbrechen, somit gar nicht keimen. Entsprechend verschiebt sich das Zahlenverhältnis nach einem Drittel sensitive und zwei Drittel resistente Pflanzen gegenüber der Selektion. Die molekular-biologischen Arbeiten bzw. Analysen wurden wie unter Beispiel 1 bis 3 beschrieben durchgeführt.

Die Linie KO-T3-02-40322-2 segregiert für eine embryoletale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-40322-2 wurde für das Restriktionsenzym Mun I mittels Adapter-PCR ein ca. 620 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 5261 des BACs MPX5, Accession AP002048 auf Chromosom 3. Durch die Insertion im Promotorbereich, ca. 243 bp stromaufwärts des Leserasters, wird die Transkription des ORF MPX5.1 verhindert und dadurch die Genfunktion ausgeschaltet. Dieser ORF kodiert für ein Protein mit Ähnlichkeit zu einem unbekannten Protein. Die

Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- 5 Die Linie KO-T3-02-40309-1 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-40309-1 wurde mittels Adapter-PCR für das Enzym Mun I ein ca. 900 bp-großes Fragment für die rechte T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 38553 des BACs F2809, Accession AL137080 auf Chromosom 3. Durch die Insertion im Promotorbereich, ca. 24 bp stromaufwärts des Leserasters, wird die Transkription des ORF F2809.140 verhindert und dadurch die Genfunktion ausgeschaltet. Dieser ORF kodiert für ein Protein mit großer Ähnlichkeit zu INT6, einem brustkrebs-assoziierten Protein und mit Ähnlichkeit zu einem „initiation-factor 3“ Protein. Die
- 10 Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.
- 15

- Die Linie KO-T3-02-40309-1 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-40309-1 wurde mittels Adapter-PCR für das Enzym Mun I ein ca. 900 bp-großes Fragment für die rechte T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 38553 des BACs F2809, Accession AL137080 auf Chromosom 3. Durch die Insertion im Promotorbereich, ca. 515 bp stromaufwärts des Leserasters, wird die Transkription des ORF F2809.150 verhindert und dadurch die Genfunktion ausgeschaltet. Dieser ORF kodiert für ein Protein mit großer Ähnlichkeit zur DNA Helikase YGL150c aus Saccharomyces. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen
- 20
- 25
- 30 Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- Die Linie KO-T4-02-00666-4 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T4-02-00666-4 wurde mittels Adapter-PCR für das Enzym Bgl II ein ca. 390 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 9358 des BACs MKN22, Accession AB019234 auf Chromosom 5. Durch die Insertion im 3'-UTR-Bereich, ca. 82 bp stromabwärts des Leserasters, wird höchstwahrscheinlich das Transkript des ORF MKN22.2 destabilisiert und dadurch die Genfunktion
- 35
- 40 ausgeschaltet. Dieser ORF kodiert für ein Protein mit Ähnlichkeit zu einem RNA-

bindenden Proteins. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- 5 Die Linie KO-T4-02-00666-4 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T4-02-00666-4 wurde mittels Adapter-PCR für das Enzym Spe I ein ca. 650 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition
- 10 48978 des BACs MEE6, Accession AB010072 auf Chromosom 5. Durch die Insertion in das offene Leseraster wird der ORF MEE6.19 unterbrochen und inaktiviert. Dieser ORF kodiert für ein Protein mit großer Ähnlichkeit zu einem unbekannten Protein. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der
- 15 erwarteten Größe lieferte, bestätigt.

- Die Linie KO-T3-02-41568-2 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-41568-2 wurde mittels Adapter-PCR für das Enzym Bgl II ein ca. 500 bp-
- 20 großes Fragment für die rechte T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 6993 des BACs T19L18, Accession AC004747 auf Chromosom 2. Durch die Insertion im 3'-UTR-Bereich, ca. 285 bp stromabwärts des Leserasters, wird höchstwahrscheinlich das Transkript des ORF At2g26150 destabilisiert und dadurch die Genfunktion
- 25 ausgeschaltet. Dieser ORF kodiert für einen putativen Hitzeschock-Transkriptionsfaktor. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- 30 Die Linie KO-T3-02-42903-1 segregiert für eine embryonale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-42903-1 wurde mittels TAIL-PCR für den degenerierten Primer ADP3 (5'-WGTGNAGWANCANAGA-3') ein ca. 1300 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte
- 35 die Insertion der T-DNA in dieser Linie an Basenpaarposition 25933 des BACs T1E2, Accession AC006929 auf Chromosom 2. Durch die Insertion in das offene Leseraster wird der ORF At2g28030 unterbrochen und inaktiviert. Dieser ORF kodiert für ein putatives, chloroplastidäres an das DNA-Nukleoid bindende Protein. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem

T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- Die Linie KO-T3-02-41395-1 segregiert für eine embryothale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-41395-1 wurde mittels Adapter-PCR für das Enzym *Mun I* ein ca. 910 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 153501 des BACs ATCHRIV25, Accession AL161513 auf Chromosom 4. Durch die Insertion in das Gen wird der ORF AT4g08990 unterbrochen und inaktiviert. Dieser ORF kodiert für ein Protein mit Ähnlichkeit zu einer putativen Met2-Typ Cytosin DNA-Methyltransferase, das starke Ähnlichkeit zu einer DNA-(cytosine-5-)-Methyltransferase aus *Arabidopsis thaliana* aufweist. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

- Die Linie KO-T3-02-44634-4 segregiert für eine embryothale Mutation, die mit dem Resistenzmarker und damit der T-DNA cosegregiert. Für die Linie KO-T3-02-44634-4 wurde mittels TAIL-PCR für den degenerierten Primer ADP8 (5'-NTGCGASWGANWAGAA-3') ein ca. 800 bp-großes Fragment für die linke T-DNA-Border identifiziert. Die Sequenzierung dieses Fragmentes offenbarte die Insertion der T-DNA in dieser Linie an Basenpaarposition 16225 des BACs F12B17, Accession AL353995 auf Chromosom 5. Durch die Insertion in das offene Leseraster wird der ORF F12B17_70 unterbrochen und inaktiviert. Dieser ORF kodiert für ein putatives Protein mit Ähnlichkeit zu einem postulierten Protein aus *Arabidopsis thaliana*. Die Insertion der T-DNA in den genannten ORF wurde mittels einer Kontroll-PCR, welche mit einem T-DNA-spezifischen und einem ORF-spezifischen Primer ein Fragment der erwarteten Größe lieferte, bestätigt.

Patentansprüche

1. Verfahren zur Identifizierung von Substanzen mit herbizider Wirkung, dadurch gekennzeichnet, dass:

5

a) die Expression oder die Aktivität des Genprodukts einer Nukleinsäure oder eines Gens umfassend:

10

aa) Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51 dargestellten Sequenz;

15

20

bb) Nukleinsäuresequenz, die sich aufgrund des degenerierten genetischen Codes aus den durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen ableiten lässt;

25

30

cc) Nukleinsäuresequenz, die ein Derivat oder ein Fragment der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 der SEQ ID NO: 51 dargestellten Nukleinsäuresequenzen ist, und mindestens 60 % Homologie auf Nukleinsäureebene aufweist;

35

- 5 dd) Kleinsäuresequenz, die für Derivate oder mente der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäuresequenzen codiert, die mindestens 50 % Homologie auf Aminosäureebene aufweisen;
- 10
- 15 ee) Nukleinsäuresequenz, die für ein Fragment oder ein Epitope eines Polypeptides codiert, das spezifisch an einem Antikörper bindet, wobei der Antikörper spezifisch an ein Polypeptid bindet, das von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 der SEQ ID NO: 51 dargestellten Sequenz codiert wird;
- 20
- 25 ff) Nukleinsäuresequenz, die für ein Fragment einer in aa) dargestellten Nukleinsäure codiert und das eine "Translation Releasing Factor"-Aktivität, eine Kobalamin-Synthese-Aktivität, Arginyl-tRNA-Synthase-Aktivität, eine RNA-Helicase-Aktivität, eine GTP-Bindeprotein-Aktivität, eine Pseudouridylat-Synthase-Aktivität, eine Adenylatkinase-Aktivität, eine Preprotein-Translocase secA Precursorprotein-Aktivität, eine DCL-Protein-Aktivität, eine Arginin-tRNA-Ligase-Aktivität, eine plastidäre Gluthathion-Reduktase-Aktivität, eine Transkriptionsfaktor Sigma-Aktivität, eine Calmodulin-Aktivität, eine INT6-Aktivität, eine Helikase YGL150c-Aktivität, eine RNA-bindende Aktivität, eine Hitzeschock-Transkriptionsfaktor-Aktivität, eine chloroplastidäre DNA-Nukleoid bindende Aktivität oder eine Met2-Typ Cytosin DNA Methyltransferase-Aktivität hat; und/oder
- 30
- 35
- 40 gg) Nukleinsäuresequenz, die für Derivate der Polypeptide mit den in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16,

5 SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24,
SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32,
SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40,
SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48,
SEQ ID NO: 50 oder SEQ ID NO: 52 dargestellten Aminosäure-
sequenzen codiert, die mindestens 20 % Homologie auf Aminosäure-
ebene aufweist und eine äquivalente biologische Aktivität besitzt;
oder

10 b) die Expression oder Aktivität einer Aminosäuresequenz, die von einer
Nukleinsäuresequenz nach aa) bis gg) codiert wird,

beeinflusst wird und solche Substanzen ausgewählt werden, die die Expression
oder die Aktivität reduzieren oder blockieren.

15

2. Verfahren nach Anspruch 1, wobei die Expression oder die Aktivität der Nuklein-
säure oder des Proteins dadurch reduziert oder blockiert wird, dass die

- 20 a) Transcription,
b) Translation,
c) Prozessierung und/oder
d) Modifikation

25

der Nukleinsäuresequenz oder Aminosäuresequenz in Anspruch 1 reduziert oder
blockiert wird.

3. Verfahren nach Anspruch 1 oder 2, wobei die Aktivität der Nukleinsäure oder des
Proteins durch eine niedermolekulare Substanz reduziert oder blockiert wird.

30 4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass
die Identifizierung der Substanzen in einem High-Throughput-Screening (HTS)
durchgeführt wird.

35 5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass
man die ausgewählten Substanzen auf eine Pflanze verbringt, um die herbizide
Aktivität der Substanzen zu testen und die Substanzen auswählt, die eine herbi-
zide Aktivität zeigen.

40 6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass das
Verfahren in einem Organismus durchgeführt wird.

7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass als Organismus Bakterien, Hefen, Pilze oder Pflanzen verwendet werden.
- 5 8. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass ein Organismus verwendet wird, der eine konditionale oder natürliche Mutante einer der in Anspruch 1 beschriebenen Sequenzen ist.
9. Nukleinsäurekonstrukt enthaltend eine Nukleinsäuresequenz, dargestellt in
10 Anspruch 1, wobei die Nukleinsäuresequenz mit einem oder mehreren Regulationssignalen verknüpft ist.
10. Substanz identifiziert nach einem Verfahren nach einem der Ansprüche 1 bis 8, wobei die Substanz ein Molekulargewicht von kleiner 1000 Dalton und größer
15 50 Dalton und einen Ki-Wert kleiner 10^{-7} M hat.
11. Substanz identifiziert nach einem Verfahren nach einem der Ansprüche 1 bis 8, wobei es sich bei der Substanz um eine proteinogene Substanz oder um eine Antisense-RNA handelt.
20
12. Substanz nach Anspruch 11, wobei es sich um einen Antikörper gegen das durch eine der in Anspruch 9 dargestellten Sequenz codiertes Protein handelt.
13. Nukleinsäurekonstrukt nach Anspruch 9, wobei im Nukleinsäurekonstrukt
25 zusätzliche weitere Nukleinsäuresequenzen enthalten sind.
14. Vektor enthaltend ein Nukleinsäurekonstrukt gemäß Anspruch 9 oder 13.
15. Organismus enthaltend mindestens ein Nukleinsäurekonstrukt gemäß
30 Anspruch 9 oder 13 oder mindestens einen Vektor gemäß Anspruch 14.
16. Organismus nach Anspruch 15, wobei es sich bei dem Organismus um eine Pflanze, einen Mikroorganismus oder ein nicht humanes Tier handelt.
- 35 17. Transgene Pflanze enthaltend ein funktionelles oder nicht funktionelles Nukleinsäurekonstrukt gemäß Anspruch 9 oder 13 oder einen Vektor gemäß Anspruch 14.
- 40 18. Verwendung eines Nukleinsäurekonstrukt gemäß Anspruch 9 oder 13 oder eines Vektors gemäß Ansprüche 14 zur Herstellung von transgenen Pflanzen.

19. Verfahren zur Identifikation eines Antagonisten von Proteinen, die durch eine Nukleinsäuresequenz gemäß Anspruch 9 oder 13 codiert werden indem folgende Verfahrensschritte durchlaufen werden
- 5
- i) Inkontaktbringen von Zellen, die das Protein exprimieren, oder des Proteins mit einem Kandidatenstoff;
 - ii) Testen der biologischen Aktivität des Protein;
 - 10 iii) Vergleichen der biologischen Aktivität des Proteins mit einer Standardaktivität in Abwesenheit des Kandidatenstoffs, wobei eine Verringerung der biologischen Aktivität des Proteins anzeigt, dass der Kandidatenstoff ein Antagonist ist.
- 15
20. Verfahren nach Anspruch 19, dadurch gekennzeichnet, dass man den nach Anspruch 19 Buchstabe iii) identifizierten Antagonisten auf eine Pflanze bringt, um seine herbizide Aktivität zu testen und die Antagonisten auswählt, die eine herbizide Aktivität zeigen.
- 20
21. Verfahren zur Bekämpfung unerwünschten Pflanzenwuchses, dadurch gekennzeichnet, dass man eine herbizid wirksame Menge einer Substanz identifiziert nach einem Verfahren gemäß Anspruch 1 bis 8 oder eines Antagonisten identifiziert nach einem Verfahren gemäß Anspruch 19 oder 20 auf Pflanzen oder deren Lebensraum oder auf Pflanzen und deren Lebensraum einwirken lässt.
- 25
22. Verwendung einer Substanz identifiziert nach einem Verfahren nach einem der Ansprüche 1 bis 8 oder eines Antagonisten identifiziert nach einem Verfahren gemäß Anspruch 19 oder 20 als Herbizid oder zur Wachstumsregulierung von Pflanzen.
- 30
23. Verfahren zur Erzeugung veränderter Genprodukte, die von den Nukleinsäuresequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, deren
- 35

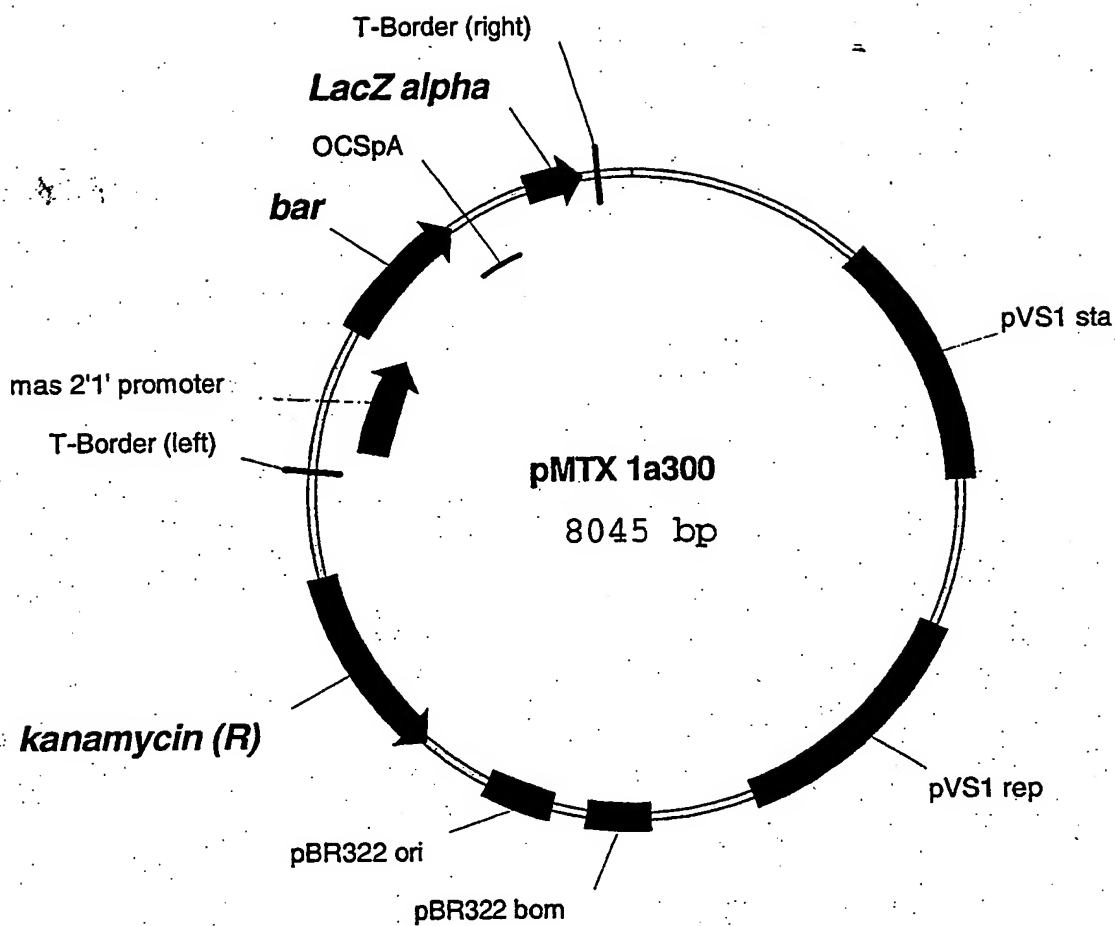
Derivate oder Fragmente gemäß Anspruch 1 codiert werden, dadurch gekennzeichnet, dass es folgende Prozessschritte umfasst:

- 5 a) Expression der von den SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49 oder SEQ ID NO: 51, deren Derivate oder Fragmente gemäß Anspruch 1 codierten Proteine in einem heterologen System oder in einem Zellfreien System
- 10 b) Randomisierte oder gerichtete Mutagenese des Proteins durch Modifikation der Nukleinsäure,
- 15 c) Messung der Interaktion des veränderten Genprodukts mit dem Herbizid oder der biologischen Aktivität des veränderten Genproduktes in Gegenwart des Herbizids,
- 20 d) Identifizierung von Derivaten des Proteins die eine geringere Interaktion aufweisen oder in ihrer Aktivität weniger beeinflusst sind,
- 25 e) Testung der biologischen Aktivität des Proteins nach Applikation des Herbizides,
- f) Auswahl der Nukleinsäuresequenzen, die oder deren Genprodukte eine veränderte biologische Aktivität gegenüber dem Herbizid aufweisen.
- 30 24. Verfahren nach Anspruch 23, dadurch gekennzeichnet, dass die gemäß Anspruch 23 f) ausgewählten Sequenzen in einen Organismus eingebracht werden.
- 35 25. Verfahren zur Erzeugung transgener Pflanzen, die gegen Substanzen gefunden nach einem Verfahren nach einem der Ansprüche 1 bis 8 oder einem Verfahren gemäß Anspruch 19 oder 20 resistent sind, dadurch gekennzeichnet, dass in diesen Pflanzen Nukleinsäuren mit den in Anspruch 1 beschriebenen Sequenzen überexprimiert werden.

26. Organismus hergestellt nach einem Verfahren gemäß Anspruch 23 oder 24 oder einem Verfahren gemäß Anspruch 25.
- 5 27. Mittel, enthaltend eine herbizid wirksame Menge mindestens einer Substanz identifiziert nach einem Verfahren nach einem der Ansprüche 1 bis 8 oder eines Antagonisten identifiziert nach einem Verfahren gemäß Anspruch 19 oder 20 und mindestens einen inerten flüssigen und/oder festen Trägerstoff sowie gegebenenfalls mindestens einen oberflächenaktiven Stoff.
- 10 28. Mittel, enthaltend eine das Wachstum regulierende Menge mindestens einer Substanz identifiziert nach einem Verfahren nach einem der Ansprüche 1 bis 8 oder eines Antagonisten identifiziert nach einem Verfahren gemäß Anspruch 19 oder 20 und mindestens einen inerten flüssigen und/oder festen Trägerstoff sowie gegebenenfalls mindestens einen oberflächenaktiven Stoff.
- 15 29. Zusammensetzung, enthaltend die Substanz nach einem der Ansprüche 10 bis 12 oder einen Antagonisten identifiziert gemäß Anspruch 19.
- 20 30. Kit, umfassend das Nukleinsäurekonstrukt nach Anspruch 9 oder 13, die Substanz nach einem der Ansprüche 10 bis 12, einen Antagonisten identifiziert gemäß Anspruch 19 oder 20 die Zusammensetzung nach Anspruch 29.

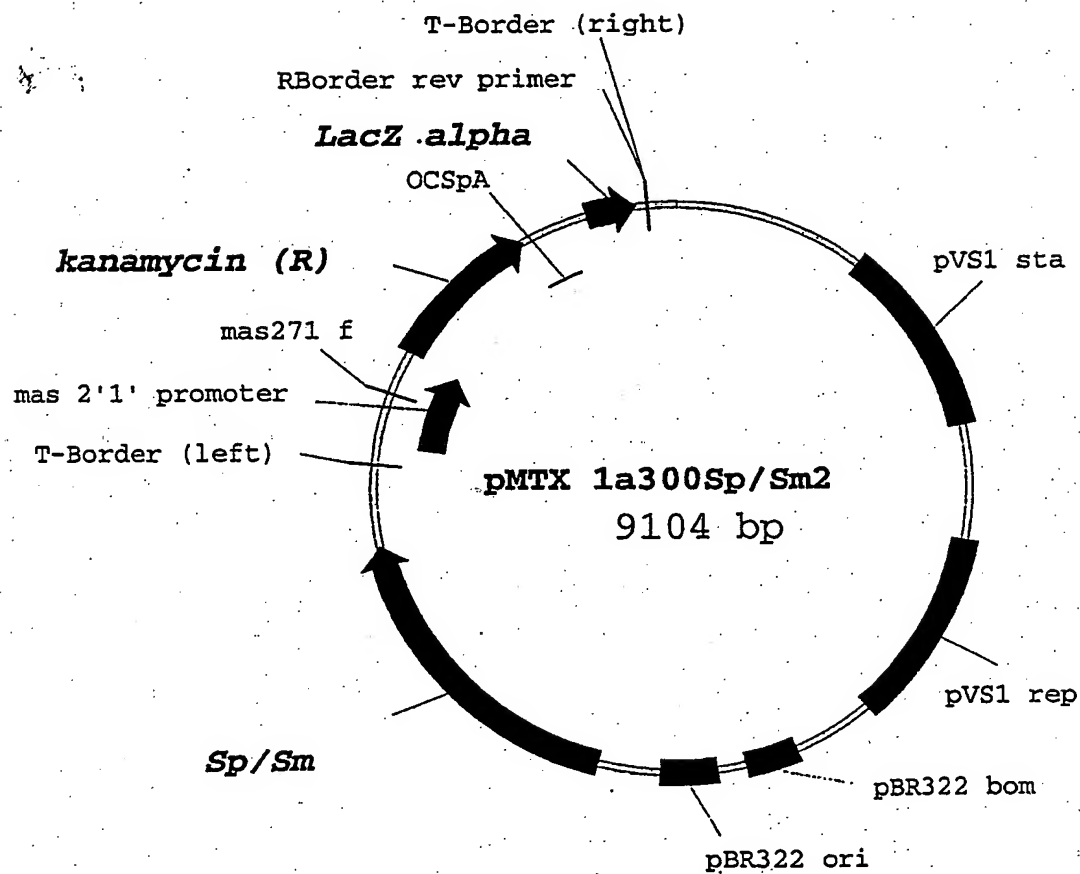


Figur 1: pMTX 1a300



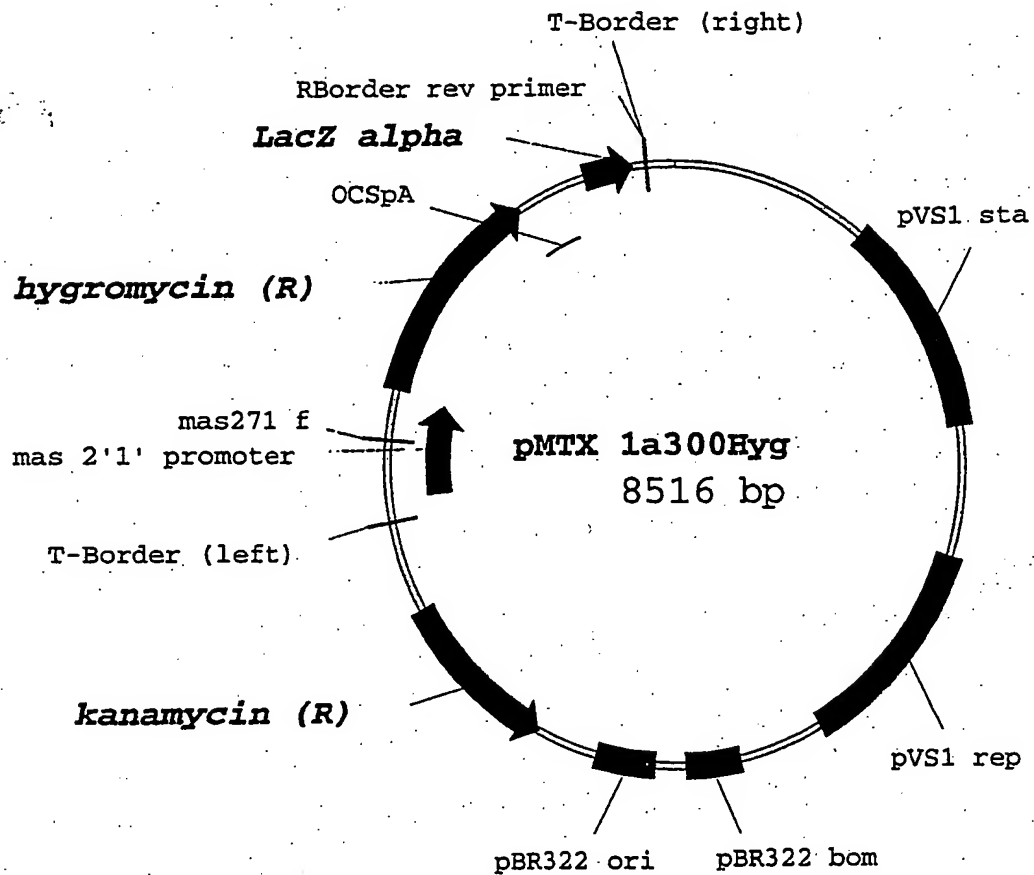


Figur 2. Vektor mit Kanamycin-Resistenz zur Selektion in Pflanzen und Spectomycin/Streptomycin-Resistenz zur Selektion in Bakterien





Figur 3. Vektor mit Hygromycin-Resistenz zur Selektion in Pflanzen und Kanamycin-Resistenz zur Selektion in Bakterien





SEQUENCE LISTING

10/524765

DT12 Rec'd PCT/PTO 16 FEB 2005

<110> Metanomics GmbH & Co. KGaA

<120> Verfahren zur Identifizierung von Substanzen mit herbizider Wirkung

<130> 53851

<150> DE 102 38 434.7

<151> 2002-08-16

<160> 52

<170> PatentIn version 3.1

<210> 1

<211> 1230

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1230)

<223>

<400> 1
 atg gcg gca aag att att ggt gga tgc tgc tca tgg cga cgc ttt tac 48
 Met Ala Ala Lys Ile Ile Gly Gly Cys Cys Ser Trp Arg Arg Phe Tyr
 1 5 10 15
 agg aag aga aca tca tct cga ttt ctg att ttc tct gtt cga gcc tct 96
 Arg Lys Arg Thr Ser Ser Arg Phe Leu Ile Phe Ser Val Arg Ala Ser
 20 25 30
 agt tcc atg gat gac atg gac acc gtc tac aag caa ttg gga ttg ttt 144
 Ser Ser Met Asp Asp Met Asp Thr Val Tyr Lys Gln Leu Gly Leu Phe
 35 40 45
 tca cta aag aag aag att aaa gat gtt gtt ctt aag gct gag atg ttt 192
 Ser Leu Lys Lys Lys Ile Lys Asp Val Val Leu Lys Ala Glu Met Phe
 50 55 60
 gca ccg gat gct ctt gag ctt gaa gaa gag cag tgg ata aag caa gaa 240
 Ala Pro Asp Ala Leu Glu Leu Glu Glu Gln Trp Ile Lys Gln Glu
 65 70 75 80
 gaa aca atg cgt tac ttt gat tta tgg gat gat ccc gct aaa tct gat 288
 Glu Thr Met Arg Tyr Phe Asp Leu Trp Asp Asp Pro Ala Lys Ser Asp
 85 90 95

gag att ctt ctc	taa gct gat cga gct aaa gca gtc gat	ctc	336
Glu Ile Leu Leu	Lys Leu Ala Asp Arg Ala Lys Ala Val Asp Ser Leu		
100	105	110	
aaa gac ctc	aaa tac aag gct gaa gaa gct aag ctg atc ata caa ttg		384
Lys Asp Leu Lys Tyr Lys Ala Glu Glu Ala Lys Leu Ile Ile Gln Leu			
115	120	125	
ggt gag atg gat gct ata gat tac agt ctc ttt gag caa gcc tat gat			432
Gly Glu Met Asp Ala Ile Asp Tyr Ser Leu Phe Glu Gln Ala Tyr Asp			
130	135	140	
tca tca ctc gat gta agt aga tcg ttg cat cac tat gag atg tct aag			480
Ser Ser Leu Asp Val Ser Arg Ser Leu His His Tyr Glu Met Ser Lys			
145	150	155	160
ctt ctt agg gat caa tat gac gct gaa ggc gct tgt atg att atc aaa			528
Leu Leu Arg Asp Gln Tyr Asp Ala Glu Gly Ala Cys Met Ile Ile Lys			
165	170	175	
tct gga tct cca ggc gca aaa tct cag gat ttg cag ata tgg aca gag			576
Ser Gly Ser Pro Gly Ala Lys Ser Gln Asp Leu Gln Ile Trp Thr Glu			
180	185	190	
caa gtt gta agt atg tat atc aaa tgg gca gaa agg cta ggc caa aac			624
Gln Val Val Ser Met Tyr Ile Lys Trp Ala Glu Arg Leu Gly Gln Asn			
195	200	205	
gcg cgg gtg gct gag aaa tgt agt tta ttg agt aat aaa agt ggc gta			672
Ala Arg Val Ala Glu Lys Cys Ser Leu Leu Ser Asn Lys Ser Gly Val			
210	215	220	
agt tca gcc acg ata gag ttt gaa ttc gag ttt gct tat ggt tat ctc			720
Ser Ser Ala Thr Ile Glu Phe Glu Phe Glu Phe Ala Tyr Gly Tyr Leu			
225	230	235	240
tta ggt gag cga ggt gtg cac cgc ctt atc ata agt tcc act tct aat			768
Leu Gly Glu Arg Gly Val His Arg Leu Ile Ile Ser Ser Thr Ser Asn			
245	250	255	
gag gaa tgt tca gcg act gtt gat atc ata cca cta ttc ttg aga gca			816
Glu Glu Cys Ser Ala Thr Val Asp Ile Ile Pro Leu Phe Leu Arg Ala			
260	265	270	
tct cct gat ttt gaa gta aag gaa ggt gat ttg att gta tcg tat cct			864
Ser Pro Asp Phe Glu Val Lys Glu Gly Asp Leu Ile Val Ser Tyr Pro			
275	280	285	
gca aaa gag gat cac aaa ata gct gag aat atg gtt tgt atc cac cat			912
Ala Lys Glu Asp His Lys Ile Ala Glu Asn Met Val Cys Ile His His			
290	295	300	
att ccg agt gga gta aca cta caa tct tca gga gaa aga aac cgg ttt			960
Ile Pro Ser Gly Val Thr Leu Gln Ser Ser Gly Glu Arg Asn Arg Phe			
305	310	315	320
gca aac agg atc aaa gct cta aac cgg ttg aag gcg aag cta ctt gtg			1008
Ala Asn Arg Ile Lys Ala Leu Asn Arg Leu Lys Ala Lys Leu Leu Val			
325	330	335	
ata gca aaa gag caa aag gtt tcg gat gta aat aaa atc gac agc aag			1056
Ile Ala Lys Glu Gln Lys Val Ser Asp Val Asn Lys Ile Asp Ser Lys			
340	345	350	
aac att ttg gaa ccg cgg gaa gaa acc agg agt tat gtc tct aag ggt			1104
Asn Ile Leu Glu Pro Arg Glu Glu Thr Arg Ser Tyr Val Ser Lys Gly			
355	360	365	
cac aag atg gtg gtt gat aga aaa acc ggt tta gag att ctg gac ctg			1152
His Lys Met Val Val Asp Arg Lys Thr Gly Leu Glu Ile Leu Asp Leu			
370	375	380	
aaa tcg gtc ttg gat gga aac att gga cca ctc ctt gga gct cat att			1200

Lys Ser Val Leu Asp Gly Asn Ile Gly Pro Leu Leu Gly Ala Ile
 385 390 395 400

agc atg aga aga tca att gat gcg att tag
 Ser Met Arg Arg Ser Ile Asp Ala Ile
 405

1230

<210> 2

<211> 409

<212> PRT

<213> Arabidopsis thaliana

<400> 2

Met Ala Ala Lys Ile Ile Gly Gly Cys Cys Ser Trp Arg Arg Phe Tyr
 1 5 10 15

Arg Lys Arg Thr Ser Ser Arg Phe Leu Ile Phe Ser Val Arg Ala Ser
 20 25 30

Ser Ser Met Asp Asp Met Asp Thr Val Tyr Lys Gln Leu Gly Leu Phe
 35 40 45

Ser Leu Lys Lys Lys Ile Lys Asp Val Val Leu Lys Ala Glu Met Phe
 50 55 60

Ala Pro Asp Ala Leu Glu Leu Glu Glu Gln Trp Ile Lys Gln Glu
 65 70 75 80

Glu Thr Met Arg Tyr Phe Asp Leu Trp Asp Asp Pro Ala Lys Ser Asp
 85 90 95

Glu Ile Leu Leu Lys Leu Ala Asp Arg Ala Lys Ala Val Asp Ser Leu
 100 105 110

Lys Asp Leu Lys Tyr Lys Ala Glu Glu Ala Lys Leu Ile Ile Gln Leu
 115 120 125

Gly Glu Met Asp Ala Ile Asp Tyr Ser Leu Phe Glu Gln Ala Tyr Asp
 130 135 140

Ser Ser Leu Asp Val Ser Arg Ser Leu His His Tyr Glu Met Ser Lys
 145 150 155 160

Leu Leu Arg Asp Gln Tyr Asp Ala Glu Gly Ala Cys Met Ile Ile Lys
 165 170 175

Ser Gly Ser Pro Gly Ala Lys Ser Gln Asp Leu Gln Ile Trp Thr Glu
 180 185 190

Gln Val Val Ser Met Tyr Ile Lys Trp Ala Glu Arg Leu Gly Gln Asn
 195 200 205

Ala Arg Val Ala Glu Lys Cys Ser Leu Leu Ser Asn Lys Ser Gly Val
 210 215 220

Ser Ser Ala Thr Ile Glu Phe Glu Phe Glu Phe Ala Tyr Gly Tyr Leu
225 230 235 240

Leu Gly Glu Arg Gly Val His Arg Leu Ile Ile Ser Ser Thr Ser Asn
245 250 255

Glu Glu Cys Ser Ala Thr Val Asp Ile Ile Pro Leu Phe Leu Arg Ala
260 265 270

Ser Pro Asp Phe Glu Val Lys Glu Gly Asp Leu Ile Val Ser Tyr Pro
275 280 285

Ala Lys Glu Asp His Lys Ile Ala Glu Asn Met Val Cys Ile His His
290 295 300

Ile Pro Ser Gly Val Thr Leu Gln Ser Ser Gly Glu Arg Asn Arg Phe
305 310 315 320

Ala Asn Arg Ile Lys Ala Leu Asn Arg Leu Lys Ala Lys Leu Leu Val
325 330 335

Ile Ala Lys Glu Gln Lys Val Ser Asp Val Asn Lys Ile Asp Ser Lys
340 345 350

Asn Ile Leu Glu Pro Arg Glu Glu Thr Arg Ser Tyr Val Ser Lys Gly
355 360 365

His Lys Met Val Val Asp Arg Lys Thr Gly Leu Glu Ile Leu Asp Leu
370 375 380

Lys Ser Val Leu Asp Gly Asn Ile Gly Pro Leu Leu Gly Ala His Ile
385 390 395 400

Ser Met Arg Arg Ser Ile Asp Ala Ile
405

<210> 3

<211> 4146

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(4146)

<223>

<400> 3

atg gct tcg ctt gtg tat tct cca ttc act cta tcc act tct aaa gca
Met Ala Ser Leu Val Tyr Ser Pro Phe Thr Leu Ser Thr Ser Lys Ala

5/114

1

10

gag cat ctc tct tcg ctc act aac agt acc aaa cat tct ttc ctc cgg Glu His Leu Ser Ser Leu Thr Asn Ser Thr Lys His Ser Phe Leu Arg 20 25 30	96
aag aaa cac aga tca acc aaa cca gcc aaa tct ttc ttc aag gtg aaa Lys Lys His Arg Ser Thr Lys Pro Ala Lys Ser Phe Phe Lys Val Lys 35 40 45	144
tct gct gta tct gga aac ggc ctc ttc aca cag acg aac ccg gag gtc Ser Ala Val Ser Gly Asn Gly Leu Phe Thr Gln Thr Asn Pro Glu Val 50 55 60	192
cgt cgt ata gtt ccg atc aag aga gac aac gtt ccg acg gtg aca atc Arg Arg Ile Val Pro Ile Lys Arg Asp Asn Val Pro Thr Val Lys Ile 65 70 75 80	240
gtc tac gtc gtc ctc gag gct cag tac cag tct tct ctc agt gaa gcc Val Tyr Val Val Leu Glu Ala Gln Tyr Gln Ser Ser Leu Ser Glu Ala 85 90 95	288
gtg caa tct ctc aac aag act tcg aga ttc gca tcc tac gaa gtg gtt Val Gln Ser Leu Asn Lys Thr Ser Arg Phe Ala Ser Tyr Glu Val Val 100 105 110	336
gga tac ttg gtc gag gag ctt aga gac aag aac act tac aac aac ttc Gly Tyr Leu Val Glu Glu Leu Arg Asp Lys Asn Thr Tyr Asn Asn Phe 115 120 125	384
tgc gaa gac ctt aaa gac gcc aac atc ttc att ggt tct ctg atc ttc Cys Glu Asp Leu Lys Asp Ala Asn Ile Phe Ile Gly Ser Leu Ile Phe 130 135 140	432
gtc gag gaa ttg gcg att aaa gtt aag gat gcg gtg gag aag gag aga Val Glu Glu Leu Ala Ile Lys Val Lys Asp Ala Val Glu Lys Glu Arg 145 150 155 160	480
gac agg atg gac gca gtt ctt gtc ttc cct tca atg cct gag gta atg Asp Arg Met Asp Ala Val Leu Val Phe Pro Ser Met Pro Glu Val Met 165 170 175	528
aga ctg aac aag ctt gga tct ttt agt atg tct caa ttg ggt cag tca Arg Leu Asn Lys Leu Gly Ser Phe Ser Met Ser Gln Leu Gly Gln Ser 180 185 190	576
aag tct ccg ttt ttc caa ctc ttc aag agg aag aaa caa gcc tct gct Lys Ser Pro Phe Phe Gln Leu Phe Lys Arg Lys Lys Gln Gly Ser Ala 195 200 205	624
ggt ttt gcc gat agt atg ttg aag ctt gtt agg act ttg cct aag gtt Gly Phe Ala Asp Ser Met Leu Lys Leu Val Arg Thr Leu Pro Lys Val 210 215 220	672
ttg aag tac tta cct agt gac aag gct caa gat gct cgt ctc tac atc Leu Lys Tyr Leu Pro Ser Asp Lys Ala Gln Asp Ala Arg Leu Tyr Ile 225 230 235 240	720
ttg agt tta cag ttt tgg ctt gga ggc tct cct gat aat ctt cag aat Leu Ser Leu Gln Phe Trp Leu Gly Gly Ser Pro Asp Asn Leu Gln Asn 245 250 255	768
ttt gtt aag atg att tct gga tct tat gtt ccg gct ttg aaa ggt gtc Phe Val Lys Met Ile Ser Gly Ser Tyr Val Pro Ala Leu Lys Gly Val 260 265 270	816
aaa atc gag tat tcg gat ccg gtt ttg ttc ttg gat act gga att tgg Lys Ile Glu Tyr Ser Asp Pro Val Leu Phe Leu Asp Thr Gly Ile Trp 275 280 285	864
cat cca ctt gct cca acc atg tac gat gat gtg aag gag tac tgg aac His Pro Leu Ala Pro Thr Met Tyr Asp Asp Val Lys Glu Tyr Trp Asn 290 295 300	912

tgg tat gac acg aag gac acc aat gac tca ctc aag agc gat	960
Trp Tyr Asp Thr Arg Asp Thr Asn Asp Ser Leu Lys Arg Lys Asp	
305 310 315 320	
gca acg gtt gtc ggt tta gtc ttg cag agg agt cac att gtg act ggt	1008
Ala Thr Val Val Gly Leu Val Leu Gln Arg Ser His Ile Val Thr Gly	
325 330 335	
gat gat agt cac tat gtg gct gtt atc atg gag ctt gag gct aga ggt	1056
Asp Asp Ser His Tyr Val Ala Val Ile Met Glu Leu Glu Ala Arg Gly	
340 345 350	
gct aag gtc gtt cct ata ttc gca gga ggg ttg gat ttc tct ggt cca	1104
Ala Lys Val Val Pro Ile Phe Ala Gly Gly Leu Asp Phe Ser Gly Pro	
355 360 365	
gta gag aaa tat ttc gta gac ccg gtg tcg aaa cag ccc atc gta aac	1152
Val Glu Lys Tyr Phe Val Asp Pro Val Ser Lys Gln Pro Ile Val Asn	
370 375 380	
tct gct gtc tcc ttg act ggt ttt gct ctt gtt ggt gga cct gca agg	1200
Ser Ala Val Ser Leu Thr Gly Phe Ala Leu Val Gly Gly Pro Ala Arg	
385 390 395 400	
cag gat cat ccc agg gct atc gaa gcc ctg aaa aag ctc gat gtt cct	1248
Gln Asp His Pro Arg Ala Ile Glu Ala Leu Lys Lys Leu Asp Val Pro	
405 410 415	
tac ctt gtg gca gta cca ctg gtg ttc cag acg aca gag gaa tgg cta	1296
Tyr Leu Val Ala Val Pro Leu Val Phe Gln Thr Thr Glu Glu Trp Leu	
420 425 430	
aac agc aca ctt ggt ctg cat ccc atc cag gtg gct ctg cag gtt gcc	1344
Asn Ser Thr Leu Gly Leu His Pro Ile Gln Val Ala Leu Gln Val Ala	
435 440 445	
ctc cct gag ctt gat gga gcg atg gag cca atc gtt ttc gct ggt cgt	1392
Leu Pro Glu Leu Asp Gly Ala Met Glu Pro Ile Val Phe Ala Gly Arg	
450 455 460	
gac cct aga aca ggg aag tca cat gct ctc cac aag aga gtg gag caa	1440
Asp Pro Arg Thr Gly Lys Ser His Ala Leu His Lys Arg Val Glu Gln	
465 470 475 480	
ctc tgc atc aga gcg att cga tgg ggt gag ctc aaa aga aaa act aag	1488
Leu Cys Ile Arg Ala Ile Arg Trp Gly Glu Leu Lys Arg Lys Thr Lys	
485 490 495	
gca gag aag aag ctg gca atc act gtt ttc agt ttc cca cct gat aaa	1536
Ala Glu Lys Lys Leu Ala Ile Thr Val Phe Ser Phe Pro Pro Asp Lys	
500 505 510	
ggt aat gta ggg act gca gct tac ctc aat gtg ttt gct tcc atc ttc	1584
Gly Asn Val Gly Thr Ala Ala Tyr Leu Asn Val Phe Ala Ser Ile Phe	
515 520 525	
tcg gtg tta aga gac ctc aag aga gat ggc tac aat gtt gaa ggc ctt	1632
Ser Val Leu Arg Asp Leu Lys Arg Asp Gly Tyr Asn Val Glu Gly Leu	
530 535 540	
cct gag aat gca gag act ctt att gaa gaa atc att cat gac aag gag	1680
Pro Glu Asn Ala Glu Thr Leu Ile Glu Glu Ile Ile His Asp Lys Glu	
545 550 555 560	
gct cag ttc agc agc cct aac ctc aat gta gct tac aaa atg gga gtc	1728
Ala Gln Phe Ser Ser Pro Asn Leu Asn Val Ala Tyr Lys Met Gly Val	
565 570 575	
cgt gag tac caa gac ctc act cct tat gca aat gcc ctg gaa gaa aac	1776
Arg Glu Tyr Gln Asp Leu Thr Pro Tyr Ala Asn Ala Leu Glu Glu Asn	
580 585 590	
tgg ggg aaa cct ccg ggg aac ctt aac tca gat gga gag aac ctt ctt	1824
Trp Gly Lys Pro Pro Gly Asn Leu Asn Ser Asp Gly Glu Asn Leu Leu	

7/114

595	600	605	
gtc tat gga aaa gcg tac ggt aat gtt ttc atc gga gtg caa cca aca Val Tyr Gly Lys Ala Tyr Gly Asn Val Phe Ile Gly Val Gln Pro Thr 610 615 620			1872
ttt ggg tat gaa ggt gat ccc atg agg ctg ctt ttc tcc aag tca gca Phe Gly Tyr Glu Gly Asp Pro Met Arg Leu Phe Ser Lys Ser Ala 625 630 635 640			1920
agt cct cat cac ggt ttt gct gct tac tac tct tat gta gaa aag atc Ser Pro His His Gly Phe Ala Ala Tyr Tyr Ser Tyr Val Glu Lys Ile 645 650 655			1968
ttc aaa gct gat gct gtt ctt cat ttt gga aca cat ggt tct ctc gag Phe Lys Ala Asp Ala Val Leu His Phe Gly Thr His Gly Ser Leu Glu 660 665 670			2016
ttt atg ccc ggg aag caa gtg gga atg agt gat gct tgt ttt ccc gac Phe Met Pro Gly Lys Gln Val Gly Met Ser Asp Ala Cys Phe Pro Asp 675 680 685			2064
agt ctt atc ggg aac att ccc aat gtc tac tat tat gca gct aac aat Ser Leu Ile Gly Asn Ile Pro Asn Val Tyr Tyr Tyr Ala Ala Asn Asn 690 695 700			2112
ccc tct gaa gct acc att gca aag agg aga agt tat gcc aac acc atc Pro Ser Glu Ala Thr Ile Ala Lys Arg Arg Ser Tyr Ala Asn Thr Ile 705 710 715 720			2160
agt tat ttg act cct cca gct gag aat gct ggt cta tac aaa ggg ctg Ser Tyr Leu Thr Pro Pro Ala Glu Asn Ala Gly Leu Tyr Lys Gly Leu 725 730 735			2208
aag cag ttg agt gag ctg ata tcg tcc tat cag tct ctg aag gac acg Lys Gln Leu Ser Glu Leu Ile Ser Ser Tyr Gln Ser Leu Lys Asp Thr 740 745 750			2256
ggg aga ggt cca cag atc gtc agt tcc atc atc agc aca gct aag caa Gly Arg Gly Pro Gln Ile Val Ser Ser Ile Ile Ser Thr Ala Lys Gln 755 760 765			2304
tgt aat ctt gat aag gat gtg gat ctt cca gat gaa ggc ttg gag ttg Cys Asn Leu Asp Lys Asp Val Asp Leu Pro Asp Glu Gly Leu Glu Leu 770 775 780			2352
tca cct aaa gac aga gat tct gtg gtt ggg aaa gtt tat tcc aag att Ser Pro Lys Asp Arg Asp Ser Val Val Gly Lys Val Tyr Ser Lys Ile 785 790 795 800			2400
atg gag att gaa tca agg ctt ttg ccg tgc ggg ctt cac gtc att gga Met Glu Ile Glu Ser Arg Leu Leu Pro Cys Gly Leu His Val Ile Gly 805 810 815			2448
gag cct cca tcc gcc atg gaa gct gtg gcc aca ctg gtc aac att gct Glu Pro Pro Ser Ala Met Glu Ala Val Ala Thr Leu Val Asn Ile Ala 820 825 830			2496
gct cta gat cgt ccg gag gat gag att tca gct ctt cct tct ata tta Ala Leu Asp Arg Pro Glu Asp Glu Ile Ser Ala Leu Pro Ser Ile Leu 835 840 845			2544
gct gag tgt gtt gga agg gag ata gag gat gtt tac aga gga agc gac Ala Glu Cys Val Gly Arg Glu Ile Glu Asp Val Tyr Arg Gly Ser Asp 850 855 860			2592
aag ggt atc ttg agc gat gta gag ctt ctc aaa gag atc act gat gcc Lys Gly Ile Leu Ser Asp Val Glu Leu Leu Lys Glu Ile Thr Asp Ala 865 870 875 880			2640
tca cgt ggc gct gtt tcc gcc ttt gtg gaa aaa aca aca aat agc aaa Ser Arg Gly Ala Val Ser Ala Phe Val Glu Lys Thr Thr Asn Ser Lys 885 890 895			2688

gga cag gtg gtc	gtg tct gac aag ctt acc tcg ctt ctt	ttt	2736
Gly Gln Val Val	Val Ser Asp Lys Leu Thr Ser Leu Leu	Phe	
900	905	910	
gga atc aat gag cca tgg gtt gag tat ttg tcc aac acc aag ttc tac	2784		
Gly Ile Asn Glu Pro Trp Val Glu Tyr Leu Ser Asn Thr Lys Phe Tyr			
915	920	925	
agg gcg aac aga gat aag ctc aga aca gtg ttt ggt ttc ctt gga gag	2832		
Arg Ala Asn Arg Asp Lys Leu Arg Thr Val Phe Gly Phe Leu Gly Glu			
930	935	940	
tgc ctg aag ttg gtg gtc atg gac aac gaa cta ggg agt cta atg caa	2880		
Cys Leu Lys Leu Val Val Met Asp Asn Glu Leu Gly Ser Leu Met Gln			
945	950	955	960
gct ttg gaa ggc aag tac gtc gag cct ggc ccc gga ggt gat ccc atc	2928		
Ala Leu Glu Gly Lys Tyr Val Glu Pro Gly Pro Gly Gly Asp Pro Ile			
965	970	975	
aga aac cca aag gtc tta cca acc ggt aaa aac atc cat gcc tta gat	2976		
Arg Asn Pro Lys Val Leu Pro Thr Gly Lys Asn Ile His Ala Leu Asp			
980	985	990	
cct cag gct att ccc aca aca gca gca atg gca agt gcc aag att gtg	3024		
Pro Gln Ala Ile Pro Thr Thr Ala Ala Met Ala Ser Ala Lys Ile Val			
995	1000	1005	
gtt gag agg ttg gta gag aga cag aag ctc gaa aac gaa ggg aaa	3069		
Val Glu Arg Leu Val Glu Arg Gln Lys Leu Glu Asn Glu Gly Lys			
1010	1015	1020	
tat ccc gag aca atc gcg ctt gtt ctt tgg gga act gac aac atc	3114		
Tyr Pro Glu Thr Ile Ala Leu Val Leu Trp Gly Thr Asp Asn Ile			
1025	1030	1035	
aaa aca tat ggg gag tct ctt ggg cag gtt ctt tgg atg att ggt	3159		
Lys Thr Tyr Gly Glu Ser Leu Gly Gln Val Leu Trp Met Ile Gly			
1040	1045	1050	
gtg aga cca att gct gat act ttt gga aga gtg aac cgt gtc gag	3204		
Val Arg Pro Ile Ala Asp Thr Phe Gly Arg Val Asn Arg Val Glu			
1055	1060	1065	
cct gtg agc tta gaa gaa cta gga agg ccg agg atc gat gta gtt	3249		
Pro Val Ser Leu Glu Glu Leu Gly Arg Pro Arg Ile Asp Val Val			
1070	1075	1080	
gtt aac tgc tca ggg gtc ttc cgt gat ctc ttt atc aac cag atg	3294		
Val Asn Cys Ser Gly Val Phe Arg Asp Leu Phe Ile Asn Gln Met			
1085	1090	1095	
aac ctt ctt gac cga gct atc aag atg gtg gcg gag cta gat gag	3339		
Asn Leu Leu Asp Arg Ala Ile Lys Met Val Ala Glu Leu Asp Glu			
1100	1105	1110	
cct gta gag caa aat ttt gta agg aaa cac gcg ttg gaa caa gca	3384		
Pro Val Glu Gln Asn Phe Val Arg Lys His Ala Leu Glu Gln Ala			
1115	1120	1125	
gag gcg ctt ggc att gat att aga gag gca gcg aca aga gtt ttc	3429		
Glu Ala Leu Gly Ile Asp Ile Arg Glu Ala Ala Thr Arg Val Phe			
1130	1135	1140	
tca aac gct tca ggg tca tac tca gcc aac atc agt ctt gct gtt	3474		
Ser Asn Ala Ser Gly Ser Tyr Ser Ala Asn Ile Ser Leu Ala Val			
1145	1150	1155	
gaa aac tcg tca tgg aac gat gag aaa cag ctt cag gac atg tac	3519		
Glu Asn Ser Ser Trp Asn Asp Glu Lys Gln Leu Gln Asp Met Tyr			
1160	1165	1170	
ttg agc cgc aaa tcg ttt gcg ttt gat agt gat gct cct gga gca	3564		
Leu Ser Arg Lys Ser Phe Ala Phe Asp Ser Asp Ala Pro Gly Ala			

1175	1180	1185	
gga atg gct gag aag aag cag gtc ttt gag atg gct ctt agc act Gly Met Ala Glu Lys Lys Gln Val Phe Glu Met Ala Leu Ser Thr 1190 1195 1200			3609
gca gaa gtc acc ttc cag aac ctg gat tct tca gag att tct ttg Ala Glu Val Thr Phe Gln Asn Leu Asp Ser Ser Glu Ile Ser Leu 1205 1210 1215			3654
act gat gtg agc cac tac ttc gat tct gac cct aca aat cta gtt Thr Asp Val Ser His Tyr Phe Asp Ser Asp Pro Thr Asn Leu Val 1220 1225 1230			3699
cag agt ttg agg aag gat aag aag aaa cca agc tct tac att gct Gln Ser Leu Arg Lys Asp Lys Lys Lys Pro Ser Ser Tyr Ile Ala 1235 1240 1245			3744
gac act aca act gca aac gcg cag gtg agg aca cta tct gag aca Asp Thr Thr Thr Ala Asn Ala Gln Val Arg Thr Leu Ser Glu Thr 1250 1255 1260			3789
gtg agg ctg gac gca aga aca aag ctg ctg aat cca aag tgg tac Val Arg Leu Asp Ala Arg Thr Lys Leu Leu Asn Pro Lys Trp Tyr 1265 1270 1275			3834
gaa gga atg atg tca agt gga tat gaa gga gtt cgt gag ata gag Glu Gly Met Met Ser Ser Gly Tyr Glu Gly Val Arg Glu Ile Glu 1280 1285 1290			3879
aag aga ctg tcc aac act gtg gga tgg agt gca acg tca ggt caa Lys Arg Leu Ser Asn Thr Val Gly Trp Ser Ala Thr Ser Gly Gln 1295 1300 1305			3924
gta gac aat tgg gtc tac gag gag gcc aac tca act ttc atc caa Val Asp Asn Trp Val Tyr Glu Glu Ala Asn Ser Thr Phe Ile Gln 1310 1315 1320			3969
gac gag gag atg ctg aac cgt ctc atg aac acc aat ccc aac tcc Asp Glu Glu Met Leu Asn Arg Leu Met Asn Thr Asn Pro Asn Ser 1325 1330 1335			4014
ttc agg aaa atg ctt cag act ttc ttg gag gcc aat ggt cgt ggc Phe Arg Lys Met Leu Gln Thr Phe Leu Glu Ala Asn Gly Arg Gly 1340 1345 1350			4059
tac tgg gac act tcc gct gaa aac ata gag aag ctc aag gaa ttg Tyr Trp Asp Thr Ser Ala Glu Asn Ile Glu Lys Leu Lys Glu Leu 1355 1360 1365			4104
tac tcg cag gtg gaa gac aag atc gaa ggg atc gat cga taa Tyr Ser Gln Val Glu Asp Lys Ile Glu Gly Ile Asp Arg 1370 1375 1380			4146

<210> 4

<211> 1381

<212> PRT

<213> Arabidopsis thaliana

<400> 4

Met	Ala	Ser	Leu	Val	Tyr	Ser	Pro	Phe	Thr	Leu	Ser	Thr	Ser	Lys	Ala
1				5					10					15	

Glu	His	Leu	Ser	Ser	Leu	Thr	Asn	Ser	Thr	Lys	His	Ser	Phe	Leu	Arg
		20					25						30		

Lys Lys His Arg Thr Lys Pro Ala Lys Ser Phe Phe Lys Lys
 35 40 45
 Ser Ala Val Ser Gly Asn Gly Leu Phe Thr Gln Thr Asn Pro Glu Val
 50 55 60
 Arg Arg Ile Val Pro Ile Lys Arg Asp Asn Val Pro Thr Val Lys Ile
 65 70 75 80
 Val Tyr Val Val Leu Glu Ala Gln Tyr Gln Ser Ser Leu Ser Glu Ala
 85 90 95
 Val Gln Ser Leu Asn Lys Thr Ser Arg Phe Ala Ser Tyr Glu Val Val
 100 105 110
 Gly Tyr Leu Val Glu Glu Leu Arg Asp Lys Asn Thr Tyr Asn Asn Phe
 115 120 125
 Cys Glu Asp Leu Lys Asp Ala Asn Ile Phe Ile Gly Ser Leu Ile Phe
 130 135 140
 Val Glu Glu Leu Ala Ile Lys Val Lys Asp Ala Val Glu Lys Glu Arg
 145 150 155 160
 Asp Arg Met Asp Ala Val Leu Val Phe Pro Ser Met Pro Glu Val Met
 165 170 175
 Arg Leu Asn Lys Leu Gly Ser Phe Ser Met Ser Gln Leu Gly Gln Ser
 180 185 190
 Lys Ser Pro Phe Phe Gln Leu Phe Lys Arg Lys Lys Gln Gly Ser Ala
 195 200 205
 Gly Phe Ala Asp Ser Met Leu Lys Leu Val Arg Thr Leu Pro Lys Val
 210 215 220
 Leu Lys Tyr Leu Pro Ser Asp Lys Ala Gln Asp Ala Arg Leu Tyr Ile
 225 230 235 240
 Leu Ser Leu Gln Phe Trp Leu Gly Gly Ser Pro Asp Asn Leu Gln Asn
 245 250 255
 Phe Val Lys Met Ile Ser Gly Ser Tyr Val Pro Ala Leu Lys Gly Val
 260 265 270
 Lys Ile Glu Tyr Ser Asp Pro Val Leu Phe Leu Asp Thr Gly Ile Trp
 275 280 285
 His Pro Leu Ala Pro Thr Met Tyr Asp Asp Val Lys Glu Tyr Trp Asn
 290 295 300
 Trp Tyr Asp Thr Arg Arg Asp Thr Asn Asp Ser Leu Lys Arg Lys Asp
 305 310 315 320

This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**